



Universidade de Aveiro Departamento de Biologia
2013

**João Luís
Oliveira Carvalho**

**O veado: análise ecológica e espacial de três
populações**



Universidade de Aveiro Departamento de Biologia
2013

**João Luís
Oliveira Carvalho**

O veado: análise ecológica e espacial de três populações

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Ecologia Aplicada, realizada sob a orientação científica do Prof. Dr. Carlos Fonseca, Professor Auxiliar com Agregação do Departamento de Biologia da Universidade de Aveiro.

o júri

presidente

Prof. Doutor João António de Almeida Seródio
Professor Auxiliar do Departamento de Biologia da Universidade de Aveiro

Doutor Pelayo Acevedo Lavandera
Investigador de Pós-Doutoramento pela Fundação para a Ciência e Tecnologia e pelo Fundo Social Europeu (arguente)

Prof. Doutor Carlos Manuel Martins Santos Fonseca
Professor Auxiliar com Agregação do Departamento de Biologia da Universidade de Aveiro (orientador)

agradecimentos

Aos meus pais pelo importante incentivo que sempre me deram. À minha irmã por sempre ter acreditado em mim apesar dos seus 11 anos;

À minha namorada Sofia pelo apoio, companhia e, sobretudo, pelo incentivo constante;

Ao meu orientador, amigo e coordenador da Unidade de Vida Selvagem do Departamento de Biologia da Universidade de Aveiro: Prof. Doutor Carlos Fonseca. Agradeço o apoio, confiança e incentivo;

Ao Prof. Doutor Alberto Gomes pelo apoio, disponibilidade e confiança;

Aos meus colegas e amigos João Santos e Rita Torres pela ajuda, sugestões, disponibilidade e revisão da dissertação;

Aos meus amigos Luciano Martins e José Silva pelos conhecimentos que me transmitiram e ideias que trocamos ao nível dos Sistemas de Informação Geográfica;

Ao Pedro Gonçalves e Luís Vidoedo pela amizade e companheirismo;

Ao Doutor Pelayo Acevedo pelos conselhos, sugestões e total disponibilidade em ajudar;

Aos diretores, técnicos e funcionários do Instituto da Conservação da Natureza e das Florestas (ICNF): Eng.^a Graça Barreira, D. Emília Martins, Sr. Agostinho Gomes, Eng.^a Anabela Possacos, Sr. Francisco Gonçalves, Sr. Manuel Fernandes, Dr. Armando Pereira, Sr. Armando Martins, D. Lúcia Gonçalves, Sr. Fernando Ribas, Eng.^o José Luís Rosa e Eng.^o Rogério Rodrigues.

palavras-chave

Machine learning, métodos de classificação, modelos de distribuição de espécies, padrões demográficos, Portugal, veado

resumo

O estudo integrado da dinâmica populacional de uma espécie exige o conhecimento detalhado do seu tamanho, estrutura e produtividade, bem como, dos fatores bióticos e abióticos que condicionam a sua ocorrência. Na Península Ibérica as populações silvestres de veado (*Cervus elaphus* L.) têm aumentado progressivamente em número e distribuição, seguindo a tendência documentada para todo o território Europeu. O valor ecológico e socioeconómico desta espécie traduz-se num crescente aumento da preocupação no que concerne à sua gestão. O presente estudo que resultou de três anos de monitorização de três populações silvestres de veado do Noroeste peninsular (Zona de Caça Nacional (ZCN) da Lombada e Reserva Regional de Caça (RRC) *Sierra de la Culebra*) e do centro de Portugal (ZCN da Lousã e Zona de Caça Turística (ZCT) da Cubeira) teve como principais objetivos: i) determinar a estrutura/composição populacional e parâmetros demográficos como a densidade, produtividade e rácio sexual; ii) caracterizar as unidades sociais e a sua variação espaço-temporal e iii) modelar a distribuição das populações avaliando a *performance* preditiva e transferibilidade dos modelos desenvolvidos. Os parâmetros populacionais são variáveis e refletem a gestão a que as populações são sujeitas. Para a população da ZCN da Lombada, num estudo anterior realizado no pico da brama, estimou-se uma densidade de aproximadamente 3,26 veados/100ha (IC95% = 2,27 – 4,70). Atualmente, o rácio macho/fêmea e cria/fêmea foram calculados em 0,78 e 0,36, respetivamente. Para a ZCN da Lousã a densidade, para o período Invernal, foi estimada em 5,52 (IC95% = 2,90 – 10,4) enquanto o rácio sexual e produtividade cifraram-se, respetivamente, nos 0,32 e 0,40. A ZCT da Cubeira, para o mesmo período, regista valores de densidade de 23,49 (IC95% = 15,83 – 34,85), um rácio sexual de 0,28 e uma produtividade de 0,42. A composição dos grupos varia ao longo do ano refletindo a etologia da espécie, enquanto o tamanho das unidades sociais regista um aumento do período outonal para o invernal. A presença do veado está positivamente relacionada com o aumento da distância a povoações e diminuição da distância a manchas florestais, matos e corpos de água. A espécie seleciona preferencialmente encostas de declives suaves expostas a sul. O modelo *random forest* foi o que apresentou uma *performance* preditiva mais elevada ($AUC = 0,864$; $HL\chi^2 = 13,8$; $p\text{-value} = 0,08$) seguido pelo *boosted regression trees* ($AUC = 0,857$; $HL\chi^2 = 11,8$; $p\text{-value} = 0,16$), *stepwise generalized linear modelling* ($AUC = 0,742$; $HL\chi^2 = 26,7$; $p\text{-value} = 0,001$) e *classification trees* ($AUC = 0,711$; $HL\chi^2 = 44,7$; $p\text{-value} = 4,19e^{-07}$).

keywords

Classification methods, demographic patterns, machine learning, Portugal, red deer, species distribution models

abstract

The integrated study of a species population dynamics requires detailed knowledge of population size, structure and productivity, but also of the biotic and abiotic factors that enhance its occurrence and distribution. In the Iberian Peninsula, red deer (*Cervus elaphus* L.) have increased both in number and distribution, following the European trend. The ecological and socioeconomic value of this species is translated in a growing concern of its effective and sustainable management. The present study is based on three years of monitoring three populations of red deer: Northwestern Iberian Peninsula (*Lombada* National Hunting Area (NHA) and *Sierra de la Culebra* Regional Hunting Reserve (RHR)) and center of Portugal (*Lousã* NHA and *Cubeira* Turistic Hunting Area (THA)) and has the following main objectives: i) determine the population structure/composition and demographic parameters (e.g. density, productivity and sex ratio); ii) characterise the social units and its spatio-temporal variation and iii) model red deer distribution, by using different statistical techniques evaluating their predictive performance and spatial transferability. The population parameters are variable and reflect the management to which the populations are subject. For the population of Northeastern Portugal specifically in *Lombada* NHA, red deer density was estimated in 3.26 animals/100ha (CI95% = 2.27 – 4.70) while sex ratio was calculated in 0.78 and productivity in 0.36. For *Lousã* NHA, red deer density was estimated in 5.52 animals/100ha (CI95% = 2.90 – 10.4). For the same population, sex ratio and productivity were calculated in 0.32 and 0.40, respectively. For *Cubeira* THA red deer density was estimated in 23.49 animals/100ha (CI95% = 15.83 – 34.85). The remaining parameters, sex ratio and productivity, were calculated in 0.28 and 0.42 respectively. In all study areas, the group composition varies along of the year, reflecting the ethology of the species, while social units' size shows an increase from autumn to winter. Red deer occurrence are positively related with the increase of the distance to villages and the distance decrease to forest patches, shrubs and water bodies. The species selects areas with smooth slopes preferentially exposed to south. Random forest model was identified as the most accurate (AUC = 0.864; $HL\chi^2$ = 13.8; p-value = 0.08) followed by boosted regression trees (AUC = 0.857; $HL\chi^2$ = 11.8; p-value = 0.16), stepwise generalized linear modelling (AUC = 0.742; $HL\chi^2$ = 26.7; p-value = 0.001) and classification trees (AUC = 0.711; $HL\chi^2$ = 44.7; p-value = $4.19e^{-07}$).

Lista de abreviaturas e siglas (Português)

EA – Época de Acasalamento

EPA – Época de pós-Acasalamento

IC – Intervalo de Confiança

PNM – Parque Natural de Montesinho

PNTI – Parque Natural do Tejo Internacional

PR – Produtividade

RNC – Reserva Nacional de Caça

RRC – Reserva Regional de Caça

RRCSC – Reserva Regional de Caça *Sierra da la Culebra*

RS – Rácio Sexual

SIC – Sítios de Importância Comunitária

ZCN – Zona de Caça Nacional

ZCNL – Zona de Caça Nacional da Lombada

ZCT – Zona de Caça Turística

ZCTC – Zona de Caça Turística da Cubeira

Lista de abreviaturas e siglas (Inglês)

AAgr – Agricultural area	GPS – Global Position System
AFor – Forest area	HL – Hosmer-Lemeshow
AHP – Analytical Hierarchic Process	HSI – Habitat Suitability Index
AIC – Akaike's Information Criterion	LNHA – Lombada National Hunting Area
Alt – Altitud/Altimetry	LR – Learning Rate
ANN – Artificial Neural Network	MANOVA – Multivariate Analysis of Variance
ASch – Shrub area	MCDM – Multi-Criteria Decision Making
AUC – Area Under the Curve	MCDS – Multiple Covariates Distance Sampling
BF – Bag Fraction	ML – Machine Learning
BRT – Boosted Regression Trees	MLC – Maximum Likelihood Classification
CART – Classification and Regression Trees	NDVI – Normalized Difference Vegetation Index
CCA – Canonical Correlation Analysis	NHA – National Hunting Area
CDS – Conventional Distance Sampling	NIR – Near Infrared
CI – Confidence Interval	NNESS – Northness
CLC – Corine Land Cover	NW – Northwest
CP – Complexity Parameter	OOB – Out-of-Bag
CT – Classification Tree	PLSR – Partial Least Square Regression
CV – Cross-Validation	RDA – Representational Difference Analysis
DAgr – Distance to agricultural fields	RF – Random Forests
DFor – Distance to forest	RMS – Root Mean Square
DRiv – Distance to main water lines	RNESS – Terrain roughness
DRoad – Distance to road network	ROC – Receiver Operating Characteristic
DSch – Distance to shrubs	SCRHR – <i>Sierra de la Culebra</i> Regional Hunting Reserve
DUrb – Distance to urban areas	SD – Standard Deviation
ENESS – Eastness	SDM – Species Distribution Models
ENFA – Ecological Niche Factor Analysis	SE – Standard Error
ETRS – European Terrestrial Reference System	Slp – Slope
EUREF – European Reference Frame	TC – Tree Complexity
GAM – Generalized Addictive Modelling	THA – Turistic Hunting Area
GARP – Genetic Algorithm for Rule set Prediction	TSA – Trend Surface Analysis
GB – Geographical Background	VIF – Variance Inflation Factors
GIS – Geographic Information System	
GLM – Generalized Linear Models	

Índice

Capítulo 1	1
Introdução.....	2
Objetivos.....	6
Estrutura da dissertação	7
Áreas de estudo.....	8
Veado – breve caracterização da espécie	15
Referências bibliográficas	17
Capítulo 2	25
Materiais e métodos.....	26
Demografia e monitorização populacional	27
Análise de dados ecológicos e ambientais	32
Referências bibliográficas	42
Capítulo 3	49
Caraterização demográfica	50
Organização social	57
Referências bibliográficas	65
Capítulo 4	68
Introduction.....	70
Material and Methods	71
Study area	71
Survey design.....	73
Environmental variables	74
Multi-criteria analysis: an overview of the AHP method.....	74
Pairwise comparisons.....	75
Data analysis	78
Results	79
Discussion	81
References	85
Capítulo 5	90
Introduction.....	92
Material and methods	94
Study area	94
Study species	96

Geographic background (GB) delimitation	96
Data collection	96
Exploratory analysis and variables selection.....	97
Modelling methods	99
Models calibration and discrimination	101
Transferability assessment	102
Results	103
Discussion	119
References	124
Capítulo 6	134
Principais resultados e conclusões gerais.....	135

Índice de Figuras

Capítulo 1

Figura 1 – Relação entre a escala temporal e espacial na seleção de <i>habitat</i> . Johnson (1980) descreve quatro níveis de seleção, i) seleção do nível geográfico de ocorrência da espécie (nível regional), ii) seleção de um domínio vital no interior do limite geográfico de ocorrência da espécie (nível populacional), iii) seleção de áreas com características gerais no interior do domínio vital da espécie (nível do domínio vital), iv) seleção de elementos particulares considerando os que estão disponíveis (nível local). Os processos que ocorrem a nível individual influenciam os processos populacionais que, por sua vez, determinam os processos evolutivos.	3
Figura 2 – Localização geográfica da Zona de Caça Nacional da Lombada e, da contígua, Reserva Regional de Caça <i>Sierra de la Culebra</i>	8
Figura 3 – Localização geográfica da Zona de Caça Nacional da Serra da Lousã.	11
Figura 4 – Localização geográfica da Zona de Caça Turística da Cubeira.	13

Capítulo 2

Figura 5 – Disposição espacial dos transectos nas áreas de estudo.	28
Figura 6 – Sequência dos principais passos adotados no desenvolvimento dos modelos apresentados no contexto da dissertação.	40

Capítulo 3

Figura 7 – Estrutura da população de veado do Noroeste da Península Ibérica. A) Dados relativos a quatro anos de amostragem na ZCN da Lombada; B) Dados relativos a dois anos de amostragem na RRC <i>Sierra de la Culebra</i> ; AB) Agrupamento dos dados das duas populações. (MJ) – Vareto; (MsubA) – Macho jovem/subadulto; (MA) – Macho adulto; (MV) – Macho velho; (Mind) – Macho indeterminado; (FJ) – Fêmea Jovem; (FA) – Fêmea adulta; (Find) – Fêmea indeterminada; (Ind) – Indeterminada.	50
Figura 8 – Estrutura da população de veado para a ZCN da Lousã (A) e para a ZCT da Cubeira (B). (MJ) – Vareto; (MsubA) – Macho jovem/subadulto; (MA) – Macho adulto; (MV) – Macho velho; (Mind) – Macho indeterminado; (FJ) – Fêmea Jovem; (FA) – Fêmea adulta; (Find) – Fêmea indeterminada; (Ind) – Indeterminada.	52
Figura 9 – Histograma que constitui a representação gráfica da função de detecção para a ZCN da Lousã ilustrando a probabilidade de detectar veados em função da distância perpendicular.	54

Figura 10 – Histograma que constitui a representação gráfica da função de detecção para a ZCT da Cubeira ilustrando a probabilidade de detectar veados em função da distância perpendicular.....	55
Figura 11 – Variação da densidade de veados nas áreas de estudo obtidas com recurso ao <i>Kernel Density Estimator</i> do software <i>ArcGIS</i> 10.1.....	56
Figura 12 – Composição dos grupos na época da brama para as três populações estudadas.....	57
Figura 13 – Distribuição do tamanho das unidades sociais para a época da brama para as três populações estudadas. A) ZCN da Lombada e RRC <i>Sierra de la Culebra</i> ; B) ZCN da Lousã e C) ZCT da Cubeira.....	59
Figura 14 – Composição dos grupos na época posterior à brama para as três populações estudadas.....	61
Figura 15 – Distribuição do tamanho das unidades sociais na época posterior à brama para as três populações estudadas. A) ZCN da Lombada e RRC <i>Sierra de la Culebra</i> ; B) ZCN da Lousã e C) ZCT da Cubeira.....	62
Capítulo 4	
Figure 16 – Location of study area.....	72
Figure 17 – Line transects in the study area.....	73
Figure 18 – Habitat suitability model output. Graduated colours from black (unsuitable areas) to white (suitable areas).....	79
Figure 19 – Area under the ROC curve used to assess the accuracy of the models: a) Analytical Hierarchic Process, b) Logistic Regression.....	80
Capítulo 5	
Figure 20 – Location of the study areas. A) Area where the models were developed. Comprises two hunting reserves: the <i>Lombada</i> National Hunting Area and the <i>Sierra de la Culebra</i> Regional Hunting Reserve. B) Location where the models were transferred: the <i>Lousã</i> National Hunting Area.....	95
Figure 21 – Trend surface analysis using a third-degree polynomial for red deer presences.....	103
Figure 22 – The decision trees (a and d) are read from the top to bottom. The nodes represent the successive splits of the training dataset defined in order to increase the homogeneity between groups. The branches indicate the path taken by individual cases and the branch width provides a visual proportion of instances that were clustered on each side. Confusion matrix is represented in figure b and the figure c illustrates the	

relative reduction of cross-validation error in relation to the size of tree and complexity parameter.	105
Figure 23 – Regression tree fitted to predict red deer abundance across the study area.	106
Figure 24 – Calibration plot showing the adjustment between predictive/estimated probabilities (x-axis) and the observed proportion of positive cases (y-axis).	106
Figure 25 – Variable importance plot for a learning rate of 0.005 (a) and 0.01 (b) showing that the distance to urban areas is the most relevant variable affecting red deer occurrence.....	107
Figure 26 – Partial dependence plots for the twelve most relevant variables identified in BRT model affecting red deer occurrence. Variables explanation and range for the modelling area (<i>Lombada-Culebra</i> hunting areas) are presented in the table 11. Rug plots at bottom of plots show distribution of records across the variable range, in deciles.	108
Figure 27 – Circular plot correlating the variables related with exposure effects. It is possible to observe an increase of presence records in the hillsides exposed to south and southwest.	109
Figure 28 – Calibration plot showing the adjustment between predictive/estimated probabilities (x-axis) and the observed proportion of positive cases (y-axis). HL ($\chi^2=11.8$ and p-value=0.16) values revealed a good adjustment between predicted and observed values.	110
Figure 29 – The error rate of the overall model decreases as each new fitted tree is added to the ensemble. Reaching 1000 trees the error rate does not decreases which indicates that the trees addition does not improve the model performance. OOB – Out-of-bag error; A – Error related with the absences represented by background data; P – Error related to the presence records.	111
Figure 30 – Cartographic representation of random forest models. The model with 5 trees represents a diffuse pattern in the occurrence probabilities, however, increasing the number of trees added to the ensemble makes this pattern more stable and defined. ...	112
Figure 31 – Graphical representation of the OOB error and AUC values. Adding trees to the ensemble reduces the error and increases discriminatory model performance reaching a plateau at 1000 trees.	113
Figure 32 – Variable importance in relation to the number of fitted trees.....	114
Figure 33 – Calibration plot showing the adjustment between predictive/estimated probabilities (x-axis) and the observed proportion of positive cases (y-axis). HL ($\chi^2=13.9$	

and p-value=0.08) values revealed a good adjustment between predicted and observed values.	115
Figure 34 – Calibration plot showing the adjustment between predictive/estimated probabilities (x-axis) and the observed proportion of positive cases (y-axis). HL ($\chi^2=26.7$ and p-value=0.001).	117
Figure 35 – Predictive probability of occurrence for red deer in the <i>Lousã</i> NHA according to the models obtained for <i>Lombada-Culebra</i> hunting areas.	118
Figure 36 – The probabilities achieved were mapped and can be interpreted as an index of habitat suitability (Aspinall & Veitch 1993). A smoother response surface provided by RF is one of the main reasons for the model high performance.	120

Índice de Tabelas

Capítulo 2

Tabela 1 – Faixas etárias para avaliação da estrutura populacional. Baseado em Marco & Gortázar (2002).30

Tabela 2 – Total de observações para as áreas de estudo por época de amostragem. EA – Época de acasalamento; EPA – Época pós-acasalamento.....34

Tabela 3 – Variáveis usadas no desenvolvimento dos modelos de classificação e regressão apresentados no capítulo 5.36

Capítulo 3

Tabela 4 – Dados relativos à densidade, rácio sexual (proporção macho/fêmea) e produtividade (proporção cria/fêmea) para as três populações. (*) Valores obtidos por Santos (2009). (**) Valores observados que poderão não corresponder ao rácio anual efetivo devido ao período em que foi realizada a amostragem. As variações no valor de densidade foram calculadas para um intervalo de confiança de 95%.53

Tabela 5 – Tamanho médio dos grupos e tamanho médio do grupo experienciado, com os respetivos intervalos de confiança a 95%, individualmente para as três populações na época da brama.....60

Tabela 6 – Tamanho médio dos grupos e tamanho médio do grupo experienciado individualmente, com os respetivos intervalos a 95%, para as três populações na época posterior à brama.....63

Capítulo 4

Table 7 – Saaty's pairwise comparisons. The values varies between 1 and 9 (factor on vertical axes is more important than the factor on horizontal axes) or 1/3 and 1/9 (factor on vertical axes is less important than the factor on horizontal axes).76

Table 8 – Comparison matrix with the relative weight assigned to the factors under analysis.....76

Table 9 – Final factor weights to be applied.77

Table 10 – Variables considered in the logistic regression model and their coefficients (β), standard deviation (SD), Wald test values and significance (ρ).....81

Capítulo 5

Table 11 – Factors and variables used to model the distribution of red deer and their range of values in *Lombada-Sierra de la Culebra* and *Lousã Mountain*. Here, are presented the average values and variation by individual unit of hexagonal grid.....98

Table 12 – Coefficients (β), standard deviation (SD) and the significance (p) for the regression model applied. Significance codes: “***” significant at 0.001, “**” significant at 0.01, “*” significant at 0.05, “.” significant at 0.1.....116

Table 13 – Correlation matrix comparing the rank order of variable selection by the four modelling techniques applied.....117

Capítulo 1

Introdução, objetivos e estrutura geral

Áreas de estudo e caracterização da espécie

Capítulo 1

Introdução

O conhecimento detalhado dos factores bióticos e abióticos que regem a dinâmica de uma população é uma condição *sine qua non* para o sucesso da sua gestão (Buckland *et al.* 1993). Associado a este aspeto, um dos maiores desafios no estudo da dinâmica das mesmas é a determinação do seu tamanho, estrutura, produtividade e tendência populacional (Mysterud *et al.* 2002; Mysterud *et al.* 2006), bem como perceber quais as variáveis ambientais, bióticas e abióticas, que determinam a sua ocorrência e distribuição (Guisan & Zimmermann 2000). Leopold (1933) refere na obra *Game Management* que a caracterização de uma população com base num critério meramente numérico é uma visão redutora. De facto, a composição, a condição e as relações entre uma população e o meio envolvente é tão ou mais importante que o seu número *per se*. O estudo da dinâmica populacional de grandes herbívoros é desenvolvido com base em dois paradigmas distintos (Gaillard *et al.* 2010): enquanto um pretende determinar como a densidade e os fatores ambientais contribuem para o crescimento populacional, o outro, tendo como base programas de monitorização a longo-prazo, descreve a dinâmica populacional partindo de valores relacionados com os respetivos parâmetros demográficos.

Nos últimos anos, em virtude das mudanças climáticas e do uso do solo, a modelação espacial tem adquirido uma elevada importância no contexto da conservação e gestão de espécies e dos seus *habitats* (ver Guisan & Theurillat 2000; Acevedo *et al.* 2011). Atualmente, os modelos de distribuição ou de adequabilidade de *habitat* são vistos como ferramentas valiosas que traduzem quantitativamente as condições e recursos ecológicos necessários à ocorrência e sobrevivência de uma população e/ou comunidade (Guisan & Thuiller 2005). De acordo com Segurado & Araújo (2004), a relação entre a ocorrência de uma espécie e um conjunto de variáveis produz dois tipos de informação: a probabilidade de uma ou mais espécies ocorrerem num local onde ainda não foram descritas e a definição de intervalos ou áreas de adequabilidade.

O conhecimento de como as espécies se distribuem no espaço e no tempo constitui um dos pilares da ecologia. As relações *habitat*-espécie processam-se a níveis hierárquicos distintos, uma vez que as espécies percecionam o *habitat* a diferentes escalas espaciais que vão desde o limite da distribuição geográfica (*Population Selection Scale*) até à seleção de determinados alimentos constituintes da sua dieta (*Local*

Selection Scale) (Johnson 1980). Assim, dependendo da escala da análise, observam-se distintos padrões de seletividade (Kie *et al.* 2002) que resultam de uma interação dinâmica dos vários fatores ambientais e da expressão complexa da ecologia e história evolutiva de uma espécie (Brown 1995) (Figura 1).

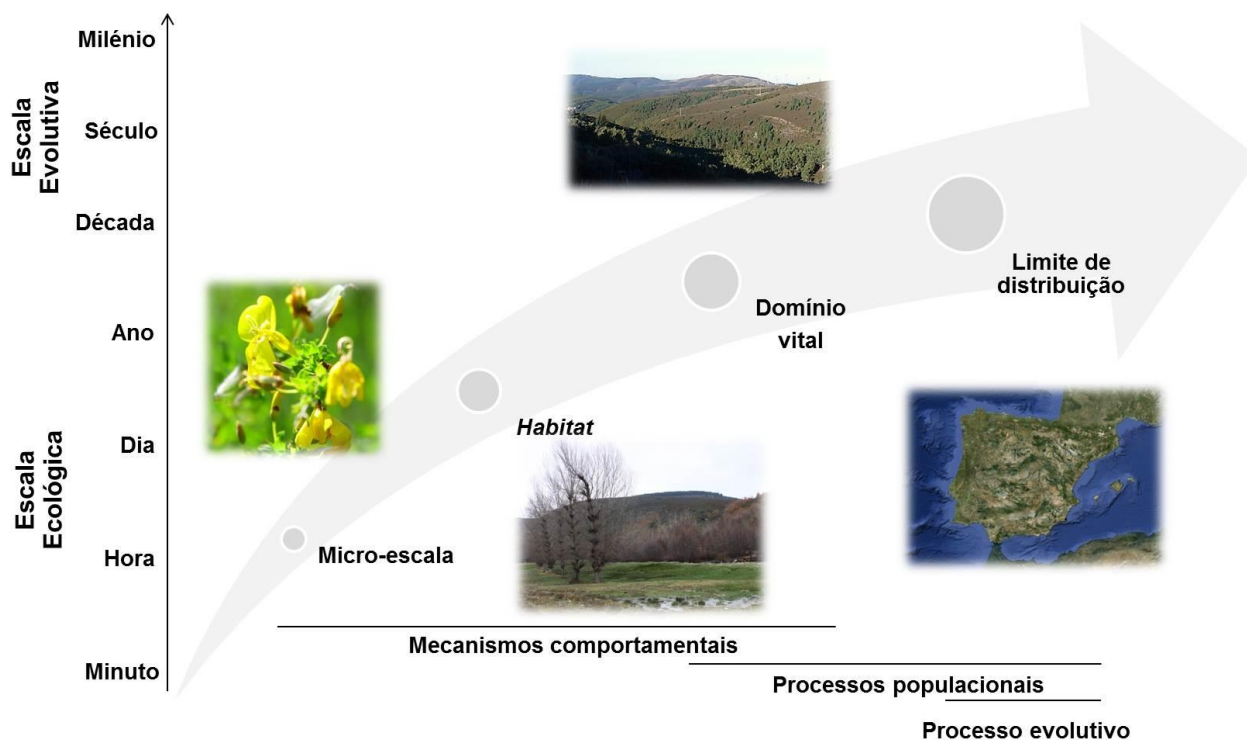


Figura 1 – Relação entre a escala temporal e espacial na seleção de *habitat*. Johnson (1980) descreve quatro níveis de seleção, **i)** seleção do nível geográfico de ocorrência da espécie (nível regional), **ii)** seleção de um domínio vital no interior do limite geográfico de ocorrência da espécie (nível populacional), **iii)** seleção de áreas com características gerais no interior do domínio vital da espécie (nível do domínio vital), **iv)** seleção de elementos particulares considerando os que estão disponíveis (nível local). Os processos que ocorrem a nível individual influenciam os processos populacionais que, por sua vez, determinam os processos evolutivos.

A atividade antrópica detém um importante papel na distribuição, tamanho populacional e estatuto de muitas espécies (Linnell & Zachos 2011), particularmente espécies de ungulados. No entanto, a complexidade das relações entre fatores naturais e antropogénicos exige uma visão holística quando o objetivo de um estudo assenta na discriminação de padrões espaço-temporais. Atualmente, existem uma plethora de técnicas de modelação que visam a exploração da correlação entre a ocorrência de uma espécie e as variáveis ambientais (Guisan & Zimmermann 2000), designadamente

árvores de regressão e classificação (Breiman *et al.* 1984), modelos lineares generalizados (McCullagh & Nelder 1989), modelos aditivos generalizados (Hastie & Tibshirani 1990), modelos mistos (McCulloch *et al.* 2008), regressões geográficas (Brunsdon *et al.* 1996), redes neuronais (Hilbert & Ostendorf 2001), correlações espaço-temporais (Epperson 2000) destacando-se, como das mais disseminadas técnicas, a Análise Factorial de Nicho Ecológico *ENFA* (Hirzel *et al.* 2002), a máxima entropia *MaxEnt* (Phillips *et al.* 2006) e o *BIOCLIM* (Beaumont *et al.* 2005). Face à variabilidade registada entre os diferentes métodos, Araújo & New (2007) sugerem que sejam testados e comparados um conjunto de modelos aumentando a fiabilidade dos resultados. De facto, são vários os estudos que avaliam a performance preditiva de diferentes técnicas de modelação (Manel *et al.* 1999; Pearce & Ferrier 2000; Thuiller *et al.* 2003; Segurado & Araújo 2004). A seleção de uma em detrimento de outra depende da espécie, do objetivo, da região e da informação disponível.

Em Portugal, ao longo das últimas três décadas, as populações de veado (*Cervus elaphus* L.) têm aumentado em número e distribuição, recuperando de uma situação de quase extinção (Vingada *et al.* 2010). Na atualidade, fruto de fenómenos de dispersão natural e reintroduções, é uma espécie comum em muitas áreas do país, particularmente nas regiões transfronteiriças (e.g. Nordeste transmontano, Parque Tejo Internacional, destacando-se também a população da serra da Lousã, este resultado de um programa de reintrodução na década de 90). A sua relevância ecológica e socioeconómica requer o seguimento contínuo das suas populações com o intuito de se obter informação relativa a parâmetros demográficos, padrões de expansão e aferir como as condicionantes ambientais e antrópicas modelam a sua distribuição. Ocorrendo em elevadas densidades, é uma espécie que exerce um significativo controlo estrutural sobre as unidades de vegetação (Schütz *et al.* 2003) revelando-se, sob determinadas circunstâncias, prejudicial para a biodiversidade (Côté *et al.* 2004; Danell *et al.* 2006; Lozano *et al.* 2007). É uma das principais presas do lobo-ibérico (*Canis lupus signatus* Cabrera, 1907), espécie estritamente protegida em Portugal (Lei nº 90/88 de 13 de Agosto) e que ostenta, segundo o Livro Vermelho dos Vertebrados de Portugal, o estatuto de conservação Em Perigo (Cabral *et al.* 2006). Juntamente com o javali (*Sus scrofa* L.) é uma espécie de caça maior de eleição, sendo a sua gestão essencial no sentido de mitigar os seus impactos no setor económico primário e nas indústrias madeireiras.

Com o presente estudo, mediante a aplicação de técnicas estatísticas e de modelação espacial, pretende-se estabelecer uma abordagem integrada na caracterização de três populações de veado (Zona de Caça Nacional da Lombada e

Reserva Regional de Caça *Sierra de la Culebra*, Zona de Caça Nacional da Serra da Lousã e Zona de Caça Turística da Cubeira) sujeitas a condições ambientais e modelos de gestão distintos. Procedeu-se a uma associação da componente demográfica e espacial tendo por base a informação recolhida ao longo de três anos de amostragem em três populações com características distintas. De facto, a gestão de populações constitui um processo complexo derivado de um conjunto de fatores intrínsecos dependentes da densidade (competição por espaço, sistema de acasalamento, predação e parasitismo) e de fatores extrínsecos independentes da mesma (temperatura, pluviosidade, *habitat*) que exercem um decisivo papel regulador (Carranza & Martínez 2002). Neste sentido, o estudo das condições que determinam o uso do espaço de uma espécie associado às questões demográficas é fundamental para o entendimento da sua ecologia.

Objetivos

“Measurement involves more than mere enumeration or census. The composition and condition of the stock is often quite as important as its numbers, and may have a bearing on the second, third, and fourth, as well as on the first step in management.”

(Aldo Leopold, 1933)

Com a presente dissertação pretende-se realizar uma análise integrada dos fatores que determinam a dinâmica de uma população silvestre, considerando a componente demográfica e identificando as variáveis que determinam a ocorrência da espécie.

Definiram-se objetivos mais específicos como forma de garantir a plena consecução do objetivo supracitado.

Demografia

- i) Caracterização da estrutura/composição populacional e determinação dos parâmetros demográficos, tais como densidade, produtividade (crias/fêmea) e rácio sexual (macho/fêmea) de três populações de veados;
- ii) Determinação dos tamanhos dos grupos e registo da sua variação espaço-temporal.

Modelação

- i) Avaliação da *performance* de um método heurístico na modelação de distribuição de espécies considerando a população de veado do Nordeste transmontano como caso de estudo;
- ii) Modelação da distribuição da população do Noroeste peninsular mediante aplicação de diferentes técnicas de classificação e regressão (*Classification and Regression Trees CART*, *Random Forests RF*, *Boosted Regression Trees BRT* e *Generalized Linear Models GLM*);
- iii) Estudo dos efeitos das variáveis (ambientais e antropogénicas) na ocorrência e abundância da espécie;
- iv) Avaliação da transferibilidade espacial dos modelos desenvolvidos.

Estrutura da dissertação

A presente dissertação encontra-se dividida em seis capítulos.

- **Capítulo 1.** Capítulo introdutório de revisão bibliográfica, com referência aos estudos e técnicas mais relevantes na temática realçando a importância de uma abordagem holística no estudo da ecologia de populações. Referência aos principais objetivos delineados no âmbito do estudo. Procede-se à caracterização e enquadramento geográfico das áreas de estudo, bem como, a uma breve descrição da espécie, o veado.
- **Capítulo 2.** Capítulo de apresentação das metodologias aplicadas. É dividido em duas componentes: demografia e modelação espacial. Na primeira é descrito o desenho experimental, a técnica de amostragem de distâncias sendo igualmente apresentados os critérios para o estudo da estrutura populacional e organização social. Na componente da modelação espacial, é realizada uma descrição sucinta dos métodos gerais e dos que serão aplicados, bem como do tipo de dados e variáveis usadas no desenvolvimento dos modelos com referência às metodologias de validação aplicadas.
- **Capítulo 3.** Capítulo onde são apresentados e discutidos os resultados relativos à demografia das três populações amostradas. Desde a estrutura populacional com referência a alguns parâmetros demográficos até à organização social. Para a organização social, os dados são apresentados na totalidade e por tipologia do grupo (machos, fêmeas ou mistos).
- **Capítulo 4.** (Artigo 1 – Publicado) Mediante a aplicação de um método heurístico, análise multicritério pelo método analítico hierárquico, procedeu-se à seleção das áreas da ZCN da Lombada com maior adequabilidade à ocorrência do veado. O modelo foi comparado com um modelo de regressão logística e ambos foram validados com recurso à curva ROC (*Receiver Operating Characteristic*).
- **Capítulo 5.** (Artigo 2 – Submetido) Aplicação de quatro métodos (*CART*, *BRT*, *RF* e *stepwise GLM*) para a determinação das variáveis mais decisivas para a ocorrência do veado. A capacidade de discriminação e calibração dos modelos foi avaliada, assim como, a sua transferabilidade espacial.
- **Capítulo 6.** Conclusões gerais, resumo e integração dos principais resultados obtidos nos três capítulos anteriores.

Áreas de estudo

Zona de Caça Nacional da Lombada (Parque Natural de Montesinho, Trás-os-Montes, Portugal) e Reserva Regional de Caça *Sierra de la Culebra* (Província de Zamora, Comunidade Autónoma de Castela e Leão, Espanha)

A Zona de Caça Nacional (ZCN) da Lombada e a Reserva Regional de Caça (RRC) *Sierra de la Culebra* localizam-se a Noroeste da região setentrional da Península Ibérica (Figura 2).

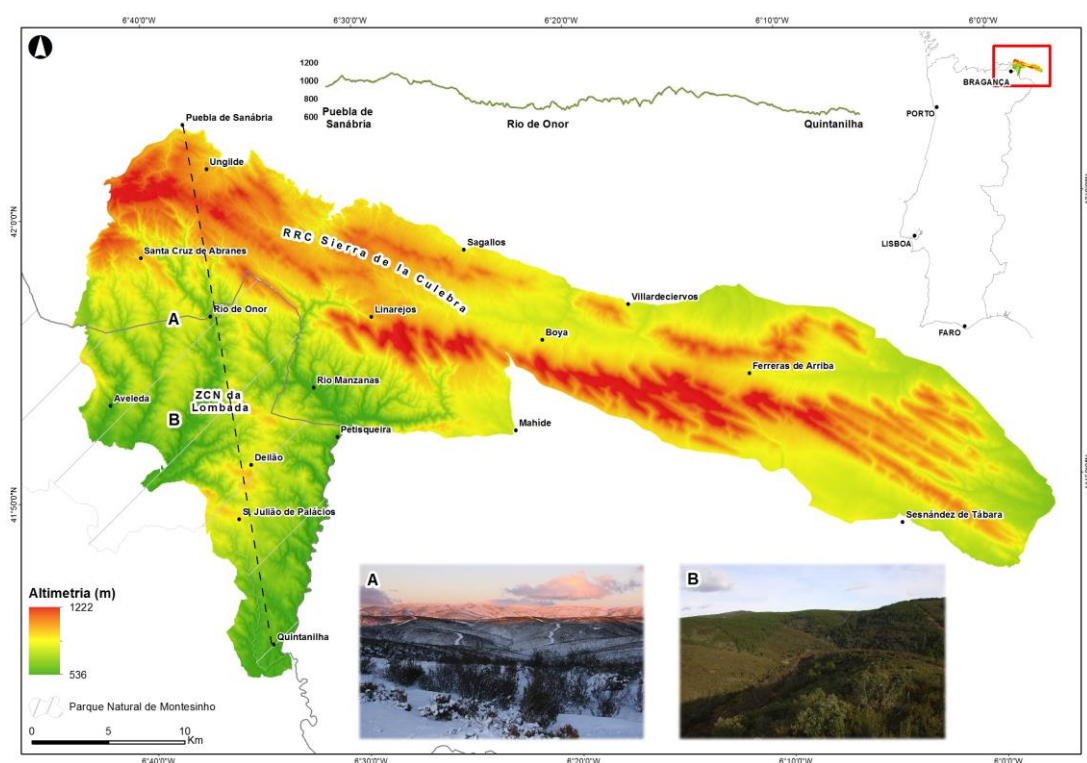


Figura 2 – Localização geográfica da Zona de Caça Nacional da Lombada e, da contígua, Reserva Regional de Caça *Sierra de la Culebra*.

A ZCN da Lombada foi criada em 1991. Da totalidade dos 20.830ha que a constituem, cerca de 90% inserem-se na área do Parque Natural de Montesinho (PNM), sendo ainda abrangidos por dois Sítios de Importância Comunitária (SIC) Rede Natura 2000: o Sítio PTCON0002 – “Montesinho/Nogueira” e a Zona de Protecção Especial para Aves Selvagens PTZPE0003 – “Serras de Montesinho e Nogueira”.

Por sua vez, em território Espanhol, na década de 70, foram constituídas 13 Reservas Nacionais de Caça (RNC), entre as quais a RNC *Sierra de la Culebra*. A Lei de

Caça de Castela de Leão estabelece que as Reservas Nacionais passem a ser denominadas de Reservas Regionais de Caça (RRC). Foi assim criada a RRC *Sierra de la Culebra* com uma área total de 67.340ha.

A existência de um núcleo populacional de veado no Nordeste de Trás-os-Montes, concelho de Bragança, está estritamente relacionada com o programa de reintrodução desta espécie na RRC *Sierra de la Culebra*, que teve lugar na década de 70. O aumento significativo dos efectivos na reserva de caça espanhola tornou possível a expansão da espécie, que começou a colonizar a área que actualmente corresponde à ZCN da Lombada, no decorrer da década de 80. Actualmente, o veado é caçado por aproximação em ambas as áreas, sendo que, na RRC *Sierra de la Culebra*, também se procede à caça seletiva. Do lado português, as quotas anuais de abate variam entre os 5–7 animais enquanto em Espanha estas atingem os 16–18 animais troféu.

Geologicamente, a região NW da Península Ibérica caracteriza-se por uma grande diversidade (Pereira *et al.* 2005). No que concerne à sua litologia, é dominada por quartzitos alternados com diferentes tipos de xistos. A erosão fluvial vertical, causada pela rede de drenagem, é dos principais processos geomorfológicos da área de estudo, sendo decisiva na formação de estreitos vales (Aguiar 2000; Santos 2006). Os intervalos altitudinais com maior expressividade na ZCN da Lombada e na RRC *Sierra de la Culebra* encontram-se, respetivamente, entre os 600 e 900 m e entre os 1000 e 1200 m (Figura 2).

Relativamente ao macroclima do Nordeste transmontano, este é moldado pelo planalto Castelhana-Leonês, a este, e a sul pelo Maciço Central conferindo-lhe características marcadamente continentais (Gonçalves 1990) com influência mediterrânica. A interioridade e morfologia desta região refletem-se em dois aspectos cruciais para a sua caracterização climatológica: precipitação e temperatura (Pereira 2006). A ZCN da Lombada localiza-se em plena Terra Fria de Planalto, com uma temperatura média que oscila entre os 10º e 12,5 °C e precipitações na ordem dos 800-900 mm anuais. A RRC *Sierra de la Culebra*, localizada mais a norte, caracteriza-se por um clima de transição entre o mediterrânico e o atlântico. Na porção este da reserva faz-se sentir, mais acentuadamente, a influência continental. De um modo geral, o intervalo de precipitação cifra-se entre os 600 e os 1000 mm.

A variabilidade geológica, climática e a heterogeneidade paisagística contribuem para um enriquecimento considerável da flora e vegetação da região. As fitocenoses têm sido igualmente moldadas pela atividade antrópica. A vegetação, embora dominada por matagais, é variada e caracterizada por bosques de carvalho-negral (*Quercus pyrenaica*), castanheiro (*Castanea sativa*), povoamentos de pinheiro-bravo (*Pinus pinaster*), pinheiro-larício (*Pinus nigra*) e pinheiro-silvestre (*Pinus sylvestris*). Nas encostas mais ensolaradas podem ser encontrados azinhais (*Quercus ilex* subsp. *rotundifolia*) mais ou menos extensos. A *Sierra de la Culebra* localiza-se nas proximidades do limite setentrional da distribuição de várias espécies mediterrânicas, nomeadamente do sobreiro (*Quercus suber*). A vegetação ripícola é maioritariamente constituída pelo amieiro (*Alnus glutinosa*), choupo-negro (*Populus nigra*), freixo (*Fraxinus angustifolia*) e salgueiro-branco (*Salix salviifolia*). As principais espécies arbustivas presentes são a carqueja (*Pterospartum tridentatum*), a esteva (*Cistus ladanifer*), a sargaça (*Halimium lasianthum* subsp. *alyssoides*) e a urze-vermelha (*Erica australis*). Para a totalidade da área (107.719ha) do SIC Montesinho/Nogueira (PTCON0002) estão descritos 24 *habitats* naturais ou semi-naturais constantes no anexo B-I do Decreto-Lei n.º 49/2005 de 24 de Fevereiro, dos quais, quatro são *habitats* prioritários (ICNB 2005). Para a RRC *Sierra de Culebra* estão identificados 19 *habitats* naturais ou semi-naturais ao abrigo do Anexo I da Diretiva 92/43/CEE.

O número de espécies animais identificadas para a região é bastante elevado, destacando-se o lobo-ibérico que exerce um importante papel regulador sobre as populações de ungulados presentes (veado, corço (*Capreolus capreolus*) e javali). Esta população lupina integra o maior núcleo da Península Ibérica. Moreira *et al.* (1997) calcularam para o Nordeste transmontano densidades de lobo que variam entre 1,6-3,1 animais/100km². Para a RRC *Sierra de la Culebra* as densidades chegam a alcançar os 10 exemplares/100km² (Gonzalez 2004). São números elevados tendo em consideração o contexto europeu, onde os valores são geralmente inferiores a 1 animal/100km² (Boitani 2000).

A agricultura, juntamente com a actividade pecuária e florestal, é a principal actividade económica da região, sendo determinante para a composição da paisagem regional e uso do solo (Aguiar 2000; Santos 2006). A densidade populacional é baixa cifrando-se, aproximadamente, nos 10 habitantes/km².

Zona de Caça Nacional da Serra da Lousã (Serra da Lousã, Beira Litoral)

A Zona de Caça Nacional (ZCN) da Serra da Lousã localiza-se na região centro de Portugal Continental, mais especificamente na região ocidental da Cordilheira Central Portuguesa (Figura 3). Criada em 2005, possui uma área total de 10.851ha abrangendo um sítio Rede Natura 2000: Sítio PTCON0060 – “Serra da Lousã”.

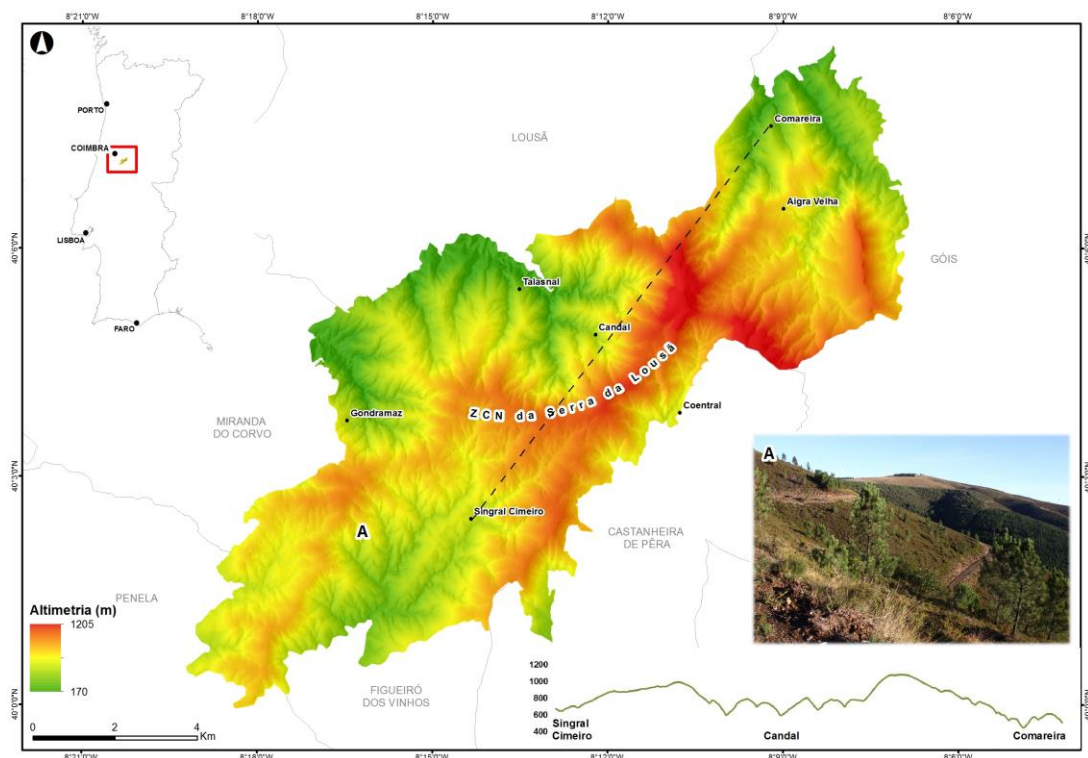


Figura 3 – Localização geográfica da Zona de Caça Nacional da Serra da Lousã.

A presença do veado na serra da Lousã resultou de um programa de reintrodução da espécie que teve início no ano de 1995. O sucesso de todo o processo traduziu-se no aumento da distribuição e densidade da espécie que é, na atualidade, alvo de exploração cinegética na serra e áreas envolventes através da modalidade de montaria.

A nível geológico, a serra da Lousã localiza-se na porção mais ocidental e central do Maciço Hespérico, local onde se observa a separação de duas zonas geotectónicas: as Zonas Centro-Ibérica e de Ossa-Morena (Pereira *et al.* 2004). A quota altitudinal varia entre os 800 e os 1200 metros, observando-se declives acentuados a norte e mais suaves a sul. Geomorfologicamente destacam-se as cristas quartzíticas, as elevações

xistentas e os depósitos de vertentes constituídos por cascalheira. Dominam os solos derivados de materiais siliciosos, maioritariamente xistos e granitos (Costa *et al.* 1998).

Na serra da Lousã fazem-se sentir influências climáticas atlânticas e mediterrânicas, o que se reflete ao nível da composição das comunidades fitocenóticas. A variabilidade climática e orográfica traduz-se na zonação do coberto vegetal. Na serra da Lousã, o coberto vegetal é bastante diferente daquele que seria o original. Atualmente existem zonas dominadas por infestantes australianas como a acácia (*Acacia* sp.) e o eucalipto (*Eucalyptus globulus*). O coberto arbustivo é representado por giestais, tojais e urzais, espécies características de terrenos pobres. Porém, subsistem bosques de folha persistente formados pelos azinhais, mais comuns nas áreas mais secas e ensolaradas, que são substituídos por bosques de caducifólias de carvalho-roble (*Quercus robur*) e castanheiro nas zonas de maior humidade. Das formações ripícolas, são os amieiros que mais se destacam pela sua representatividade e estatuto. Para a área (15.158ha) do SIC Serra da Lousã (PTCON0060) estão identificados 16 *habitats* naturais ou semi-naturais constantes no anexo B-I do Decreto-Lei n.º 49/2005 de 24 de Fevereiro, dos quais, três são *habitats* prioritários (ICNB 2005).

Além do veado, estão presentes o javali e o corço. A serra da Lousã é uma área particularmente importante para a conservação de espécies como o lagarto-de-água (*Lacerta schreiberi*) e a salamandra-lusitânica (*Chioglossa lusitanica*).

Apesar da grande maioria do sítio ser de ocupação florestal, nos vales aluvionares desenvolve-se uma atividade agrícola dominada por culturas arvenses. Relativamente à pecuária destaca-se o gado ovino e caprino. A densidade populacional de 29,2 habitantes/km² é a mais elevada para as três áreas de estudo.

Zona de Caça Turística da Cubeira (Parque Natural do Tejo Internacional, Beira Baixa)

A Zona de Caça Turística (ZCT) da Cubeira, criada em 1994, localiza-se na Beira Interior estando circunscrita ao Parque Natural do Tejo Internacional (PNTI) (Figura 4).

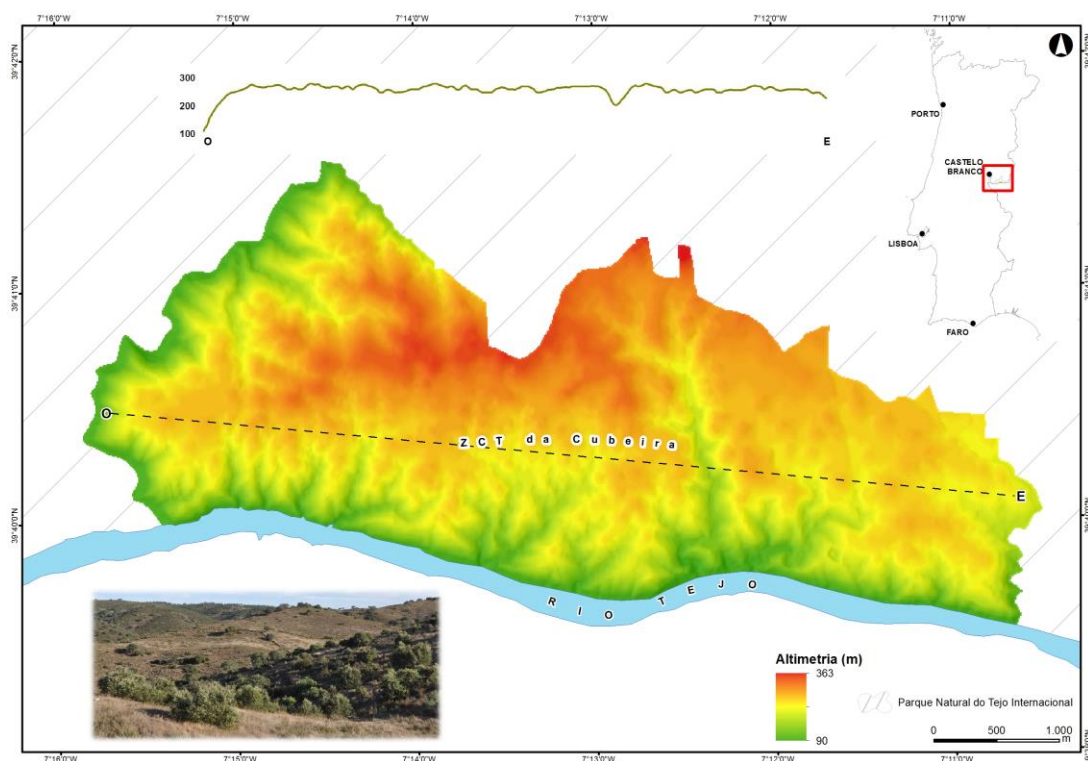


Figura 4 – Localização geográfica da Zona de Caça Turística da Cubeira.

A região correspondente à Beira Baixa integra uma área de transição entre a Cordilheira Central e a peneplanície Alentejana. Litologicamente é dominada pelo complexo xisto-grauváquico e respetivas séries metamórficas destacando-se, igualmente, alguns afloramentos graníticos. A quota altitudinal varia entre os 100 e os 300 metros. A rede hidrográfica modela o relevo das áreas xistosas que adquire um perfil mais acidentado, contrastando com o relevo suave dos terrenos detríticos. Os principais rios da região (Aravil, Erges, Ponsul, Tejo) caracterizam-se pelas suas encostas abruptas e vales profundos, sendo que o Tejo apresenta, do lado espanhol, margens suaves (ICN 2006).

Do ponto de vista climático, a área de estudo, localizando-se na província Luso-Extremadurense, caracteriza-se pelos invernos temperados e um ombroclima de sub-húmido a seco (Costa *et al.* 1998). A precipitação total anual dificilmente ultrapassa os 600 mm e concentra-se entre os meses de Outubro e Março.

A vegetação climácica do PNTI inclui as séries lusoextremadurense e bética sub-húmida do sobreiro, e mediterrânica lusoextremadurense silicícola da azinheira. Espécies como o carrasco (*Quercus coccifera*), o lentisco (*Phillyrea angustifolia*), o aderno-de-folhas-largas (*Phillyrea latifolia*) e o medronheiro (*Arbutus unedo*) são comuns. Destacam-se igualmente os freixiais, salgueirais, amiais e tamujais. A nível arbustivo, são os estevais, os giestais, os rosmaninhais, os tojais e os zambujais as formações vegetais com maior representação. Na área do PNTI (26.484ha) ocorrem 16 habitats naturais ou semi-naturais constantes no anexo B-I do Decreto-Lei n.º 49/2005 de 24 de Fevereiro, dos quais, três são habitats prioritários (ICNB 2005). O elenco florístico regista uma significativa componente endémica (Carvalhinho 2003).

A nível das comunidades faunísticas, além do veado e javali, que detêm um importante valor cinegético, destacam-se várias espécies de avifauna pelo valor conservacionista que ostentam. Destacam-se a águia-de-Bonelli (*Aquila fasciata*), a águia-imperial-ibérica (*Aquila adalberti*), a águia-real (*Aquila chrysaetos*), o abutre-negro (*Aegypius monachus*), a cegonha-negra (*Ciconia nigra*), entre outras.

A densidade populacional é muito baixa (aprox. 4 habitantes/km²). A agricultura é a atividade económica de maior expressão. Esta processa-se em diferentes sistemas, desde policultura intensiva, sistemas especializados na produção agro-florestal e sistemas agro-silvo-pastoris extensivos.

Veado – breve caracterização da espécie

O veado, maior mamífero terrestre da nossa fauna, apresenta um dimorfismo sexual considerável. O tamanho e peso corporal são muito mais variáveis nos machos do que nas fêmeas (Carranza 2008), estando o crescimento corporal e desenvolvimento da armação amplamente relacionados com as condições ambientais e densidade populacional.

É uma espécie adaptada a climas temperados. Segundo Blanco (1998), a ausência de água e a neve invernal, principalmente quando esta atinge vários centímetros de altura, são factores limitantes. É um animal que ocorre em zonas de transição (ecótonos) entre bosque e clareiras, local preferencialmente seleccionado pelos machos para marcação territorial (Ramos *et al.* 2006). No período estival selecciona encostas orientadas a norte e lugares sombrios e húmidos, enquanto no Inverno procura encostas expostas a sul e áreas descobertas e/ou ensolaradas (Guerrero & Aparicio 1997). O uso do espaço e os movimentos diários ou sazonais do veado são característicos da sua biologia, sendo influenciados pelo regime de chuvas, disponibilidade de alimento, distribuição dos pontos de água, predação, competição, presença/ausência de gado doméstico e intensidade de caça.

As áreas vitais da espécie para a Península Ibérica são menores do que as registadas para o centro e Norte da Europa. No Parque Natural de Monfragüe, registaram-se áreas de 250ha para as fêmeas e 655ha para os machos (Carranza *et al.* 1991). Na Sierra Morena obtiveram-se valores de 417ha para as fêmeas e 1.185ha para os machos (Soringuer *et al.* 1994), enquanto no Parque Nacional de Doñana, os valores para fêmeas e machos foram de 240 e 1.050ha, respetivamente (Sánchez-Prieto & Carranza 2003).

Entre setembro e finais de outubro ocorre a época da brama (cio). É um período de grande desgaste em que os machos estão constantemente ativos, o que resulta numa perda significativa da massa corporal (Barroso & Rosa 1999). À semelhança das restantes espécies de caça maior, o veado é poliétrico estacional (Guerrero & Aparicio 1997) e, como animal poliginico, os machos tentam copular com o maior número de fêmeas possível. O rácio sexual, a densidade populacional e a concentração do alimento determinam o número de fêmeas por harém (Carranza 2008). Em condições de escassez alimentar, o requisito para a selecção de uma área por parte da fêmea é a existência de recursos alimentares, em detrimento de um macho em particular. Se esta situação não se verificar, a selecção do macho é efectuada de acordo com a sua condição corporal e

tamanho ou estrutura das hastes (Bartos & Bahboub 2006). Durante a época de escassez é comum a disponibilização de suplemento alimentar, principalmente, em áreas cercadas. A agregação espacial dos animais e, em particular das fêmeas, afecta os padrões naturais de emparelhamento (Carranza & Martínez 2002; Pérez-González *et al.* 2010).

Apesar de se alimentar preferencialmente de herbáceas, sobretudo entre finais de outono e inícios do verão, o veado está adaptado a alimentar-se também de espécies lenhosas (Krojerová-Prokešová *et al.* 2010). Em regiões mediterrânicas com verões secos, as plantas lenhosas assumem uma importância acrescida na dieta da espécie (Bugalho & Milne 2003). Assim, a região geográfica, à semelhança de outros aspectos como a época do ano, o sexo e a classe etária determinam a proporção de herbáceas e lenhosas consumidas (Blanco 1998; Bugalho *et al.* 2001; Carranza 2008). Entre as espécies seleccionadas destacam-se as gramíneas e espécies arbustivas e arbóreas como a carqueja, os carvalhos, a esteva, o lentisco-bastardo (*Phillyrea angustifolia* L.), o medronheiro, a oliveira (*Olea europea* L.) a sargaça, a silva (*Rubus ulmifolius* Schott) e as urzes (Alvarez & Ramos 1991; Ferreira 1998; Orueta *et al.* 1998).

Machos e fêmeas associam-se, durante a maior parte do ano, em grupos unissexuais, no caso dos primeiros, e familiares constituídos por fêmeas e jovens dos dois sexos (Blanco 1998). A época do cio ou da brama constitui a única exceção. Uma unidade social é constituída por uma fêmea adulta e seus descendentes dos últimos dois ou três anos. As fêmeas jovens podem permanecer com a progenitora até ao primeiro parto, altura em que formam a sua unidade familiar numa área que se poderá sobrepor à da progenitora (Carranza 2002). Num grupo familiar matriarcal, as fêmeas adultas dispõem-se perifericamente enquanto as crias se localizam mais centralmente (Guerrero & Aparicio 1997). O tamanho de um grupo está relacionado com a densidade populacional e/ou época do ano (Blanco 1998). No verão utilizam áreas de maior cobertura para se alimentarem, daí os grupos serem menores. Por outro lado, no Inverno dão preferência a áreas abertas o que se repercute no aumento do tamanho dos mesmos. Grupos maiores aumentam a efetividade da proteção contra predadores.

O veado compete por recursos com várias outras espécies de herbívoros, tanto domésticos como silvestres. No caso dos domésticos, verifica-se uma competição mais intensa entre o veado e o gado caprino e/ou ovino. Vários estudos indicam que o veado além de influenciar a utilização do habitat por parte do corço é prejudicial às suas populações a partir de determinada densidade/abundância (Latham *et al.* 1996; San-José *et al.* 1997; Torres *et al.* 2012).

Referências bibliográficas

- Acevedo P., Farfán M.A., Márquez A.L., Delibes-Mateos, M., Real R. & Vargas J.M. (2011). Past, present and future of wild ungulates in relation to changes in land use. *Landscape Ecology*. 16: 19–31.
- Aguiar C. (2000). Flora e vegetação da serra da Nogueira e do Parque Natural de Montesinho. Tese de Doutoramento em Engenharia Agronómica. Universidade Técnica de Lisboa. Instituto Superior de Agronomia. Lisboa.
- Alvarez G. & Ramos J. (1991). Estrategias alimentarias del ciervo (*Cervus elaphus* L.) en Montes de Toledo. *Doñana Acta Vertebrata*. 18: 63–99.
- Araújo M.B. & New M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*. 22: 42–47.
- Barroso I. & Rosa J.L. (1999). O Veado. Coleção Património Natural Transmontano. João Azevedo Editor. Mirandela. 74pp.
- Bartos L. & Bahboub R. (2006). Antler size and fluctuating asymmetry in red deer (*Cervus elaphus*) stags and probability of becoming a harem holder in rut. *Biological Journal of Linnean Society*. 87: 59–68.
- Blanco J.C. (eds.) (1998). Mamíferos de España II. Cetáceos, Artiodáctilos, Roedores y Lagomorfos de la Península Ibérica, Baleares y Canarias. Editorial Planeta S.A. Barcelona. 416pp.
- Beaumont L.J., Hughes L. & Poulsen M. (2005). Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling*. 186: 250–269.
- Boitani L. (2000). Action Plan for the Conservation of Wolves (*Canis lupus*) in Europe. Nature and Environment, Council of Europe Publishing. 84pp.

Breiman L., Friedman J.H., Olshen R.A. & Stone C.J. (1984). Classification and regression trees. Wadsworth International Group. Belmont, CA. 358pp.

Brown J.H. (1995). Macroecology. University of Chicago Press. Chicago. 284pp.

Brunsdon C.A., Fotheringham A.S., Charlton M.E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*. 28: 281–298.

Buckland S.T., Ahmadi S., Staines W., Gordon I.J. & Youngson R.W. (1993). Estimating the minimum population size that allows a given annual number of mature red deer stags to be culled sustainably. *Journal of Applied Ecology*. 33: 118–190.

Bugalho M.N., Milne J.A. & Racey P.A. (2001). The foraging ecology of red deer (*Cervus elaphus*) in a Mediterranean environment: is a larger body size advantageous? *Journal of Zoology*. 255: 285–289.

Bugalho M.N. & Milne J.A. (2003). The composition of the diet of red deer (*Cervus elaphus*) in a Mediterranean environment: a case of summer nutritional constraint? *Forest Ecology and Management*. 181: 23–29.

Cabral M.J., Almeida J., Almeida P.R., Dellinger T., Ferrand de Almeida N., Oliveira M.E., Palmeirim J.M., Queiroz A.I., Rogado L. & Santos-Reis M. (eds.) (2006). Livro Vermelho dos Vertebrados de Portugal. 2ªEdição. Lisboa. Instituto da Conservação da Natureza. 659pp.

Carranza J., Hidalgo de Trucios S.J., Medina R., Valencia J. & Delgado J. (1991). Space use by red deer in a Mediterranean ecosystem as determined by radio-tracking. *Applied Animal Behaviour Science*. 30: 363–371.

Carranza J. (2002). Environmental effects on the evolution of mating systems in endotherms. Pp. 106-139. Em: Apollonio M., Festa-Bianchet M. & Mainardi D. (eds.). *Vertebrate Mating Systems*. World Scientific. 340pp.

Carranza J. & Martínez J. (2002) Consideraciones evolutivas en la gestión de espécies cinegéticas. Pp. 373-387. Em: Soler M. (eds.). Evolución, la base de la Biología. Proyecto Sur Ediciones. Granada. 559pp.

Carranza J. (2008). Ciervo –*Cervus elaphus* Linnaeus, 1758. Em: Carrascal L.M. & Salvador A. (eds.). Enciclopedia Virtual de los Vertebrados Españoles. Museo Nacional de Ciencias Naturales. Madrid. <http://www.vertebradosibericos.org/>.

Carvalhinho J.N.M. (2003). A Flora e a Vegetação dos *Habitats* Naturais do Parque Natural do Tejo Internacional. Volume I (Flora). Castelo Branco. Parque Natural do Tejo Internacional/Instituto da Conservação da Natureza. 78pp.

Costa J.C., Aguiar C., Capelo J.H., Lousã M., Neto C. (1998). Biogeografia de Portugal Continental. Quercetea. 0: 5–56.

Côté S.D., Rooney T.P., Tremblay J.P., Dussault C. & Waller D.M. (2004). Ecological impacts of deer over-abundance. Annual Review of Ecology Evolution and Systematics. 35: 113–147.

Danell K., Duncan P., Bergström R. & Pastor J. (2006). Large Herbivore Ecology, Ecosystem Dynamics and Conservation. Cambridge University Press. 522pp.

Epperson B.K. (2000). Spatial and space-time correlations in ecological models. Ecological Modelling. 132: 63–76.

Ferreira S. (1998). Estudo da dieta de duas populações de veado (*Cervus elaphus* Linnaeus, 1758) em Portugal. Relatório de Estágio para obtenção da licenciatura em Biologia. Universidade de Coimbra.

Gaillard J.M., Coulson T. & Festa-Bianchet M. (2010). Demographic processes: lessons from long-term, individual-based studies. Pp.98–116. Em: Owen-Smith N. (eds.). Dynamics of Large Herbivore Populations in Changing Environments. Blackwell Publishing. 217pp.

Gonçalves D.A. (1990). O uso do solo e a construção das paisagens rurais. O caso do interior de Trás-os-Montes. 1ª Jornadas sobre o Mundo Rural. Instituto Politécnico de Bragança.

Gonzales J.L.V. (2004). La gestión del lobo ibérico (*Canis Lupus signatus* Cabrera, 1907), en una población “estable”. Pp.221–240. Em: Lucio A.J. & Buruaga M.S. (eds.). Aportaciones a la Gestión Sostenible de la Caza. FEDENCA - Escuela Española de Caza. 312pp.

Guerrero Y.M. & Aparicio M.M. (1997). Bases biológicas y gestión de espécies cinegéticas en Andalucía. Servicio de Publicaciones de la Universidad de Córdoba. 182pp.

Guisan A. & Theurillat J.P. (2000). Assessing alpine plant vulnerability to climate change: a modelling perspective. *Integrated Assessment*. 1: 307–320.

Guisan A. & Zimmermann N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*. 135: 147–186.

Guisan A. & Thuiller W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*. 8: 993–1009.

Hastie T.J. & Tibshirani R. (1990). *Generalized Additive Models*. Chapman and Hall. 352pp.

Hilbert D.W. & Ostendorf B. (2001). The utility of artificial neural networks for modelling the distribution of vegetation in past, present and future climates. *Ecological Modelling*. 146: 311–327.

Hirzel A.H., Hausser J., Chessel D. & Perrin N. (2002). Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? *Ecology*. 83: 2027–2036.

ICNB (2005). Plano Sectorial Rede Natura 2000. Caracterização dos Valores Naturais. Disponível em: http://www.icn.pt/psrn2000/caract_habitat.htm.

Johnson D.H. (1980). The comparison of usage and availability measurements for evaluating resource preference. *Ecology*. 61: 65–71.

Kie J.G., Bowyer R.T., Nicholson M.C., Boroski B.B., Loft E.R. (2002). Landscape heterogeneity at differing scales: effects on spatial distribution of mule deer. *Ecology*. 83(2): 530–544.

Krojerová-Prokešová J., Barančeková M., Šustr, P. & Heurich M. (2010). Feeding patterns of red deer *Cervus elaphus* along an altitudinal gradient in the Bohemian Forest: effect of habitat and season. *Wildlife Biology*. 16: 173–184.

Latham J., Staines B.W., Gorman M.L. (1996). The relative densities of red (*Cervus elaphus*) and roe (*Capreolus capreolus*) deer and their relationship in Scottish forests. *Journal of Zoology*. 240: 285–299.

Leopold A. (1933). *Game Management*. University of Wisconsin Press. 521pp.

Linnell J.D.C. & Zachos F.E. (2011). Status and distribution patterns of European ungulates: genetics, population, history and conservation. Pp.12–53. Em: Putman R., Apollonio M. & Andersen R. (eds.). *Ungulate Management in Europe: Problems and Practices*. Cambridge University Press. 410pp.

Lozano J., Virgós E., Cabezas-Díaz S. & Mangas J.G. (2007). Increase of large game species in Mediterranean areas: Is the European wildcat facing a new threat? *Biological Conservation*. 138: 321–329.

Manel D., Dias J.M., Buckton S.T. & Ormerod S.J. (1999) Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology*. 36: 734–747.

McCullagh P. & Nelder J.A. (1989). *Generalized Linear Models*. London: Chapman and Hall. 511pp.

McCulloch C.E., Searle S.R. & Neuhaus J.M. (2008). *Generalized, Linear, and Mixed Models*. John Wiley & Sons, Inc. 424pp.

Moreira L.M., Rosa J.L., Lourenço J., Barroso I. & Pimenta V. (1997). Projecto Lobo. Relatório de Progressão 1996 (Cofinanciado pela U. E. - Programa LIFE). Parque Natural de Montesinho, Bragança, Portugal. 61pp.

Mysterud A., Coulson T. & Stenseth N.C. (2002). The role of males in the dynamics of ungulate populations. *Journal of Animal Ecology*. 71: 907–915

Mysterud A., Meisingset E.L., Veiberg V., Langvatn R., Solberg E.J., Loe L.E. & Stenseth N.C. (2006). Monitoring population size of red deer *Cervus elaphus*: an evaluation of two types of census data from Norway. *Wildlife Biology*. 13: 285–298.

Orueta J.F., Aranda Y. & García F.J. (1998). Efecto del ramoneo del ciervo (*Cervus elaphus*) sobre dos especies del matorral mediterráneo en los Montes de Toledo (Centro de España). *Galemys*. 10: 27–36.

Pearce J. & Ferrier S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*. 133: 225–245.

Pereira L.C., Sequeira A.J.D. & Gomes E.M.C. (2004). A deformação varisca do Maciço Hespérico na região da Serra da Lousã (Portugal central). *Caderno do Laboratorio Xeológico de Laje*. 29: 203–214.

Pereira P., Pereira D.I. & Alves I.C. (2005). Particularidades da morfologia granítica da Serra de Gamoneda-Montesinho (Espanha-Portugal). *Actas do X Colóquio Ibérico de Geografia*. Universidade de Évora.

Pereira, P. (2006). Património geomorfológico: conceptualização, avaliação e divulgação. Aplicação ao Parque Natural de Montesinho. Tese de Doutoramento em Ciências. Área de conhecimento em Geologia. Universidade do Minho.

Pérez-González J., Barbosa A.M., Carranza J. & Torres-Porras J. (2010). Efecto relativo de la suplementación de alimento y los recursos naturales en la distribución de hembras de ciervo en un ecosistema mediterráneo. *X Congreso Iberoamericano de Etología*. Ciudad Real.

- Phillips S.J., Anderson R.P. & Schapire R.E. (2006). Maximum entropy modelling of species geographic distributions. *Ecological Modelling*. 190: 231–259.
- Ramos J.A., Bugalho M.N., Cortez P. & Iason G.R. (2006). Selection of trees for rubbing by red and roe deer in forest plantations. *Forest Ecology and Management*. 222: 39–45.
- Sánchez-Prieto C.B. & Carranza J. (2003). Estrategias de apareamiento y uso del espacio en el ciervo ibérico. VI Jornadas de la Sociedad Española para la Conservación y Estudio de los Mamíferos. Ciudad Real.
- San-José C., Braza F., Aragón S. & Delibes J.R. (1997). Habitat use by roe and red deer in southern Spain. *Miscellanea Zoologica*. 20: 27–38.
- Santos J.M.R. (2006). La transformación del paisaje vegetal y los aprovechamientos forestales tradicionales en la sierra de la Culebra (Zamora). *Investigaciones Geográficas*. 40: 55–72.
- Schütz M., Risch A.C. Leuzinger E. Krüsi, B.O. & Achermann G. (2003). Impact of herbivory by red deer (*Cervus elaphus* L.) on patterns and processes in subalpine grasslands in the Swiss National Park. *Forest Ecology and Management*. 181: 177–188.
- Segurado P. & Araújo M.B. (2004). An evaluation of methods for modelling species distribution. *Journal of Biogeography*. 31: 1555–1568.
- Soriguer R.C., Fandos P., Bernaldez E. & Delibes J.R. (1994). El Ciervo en Andalucía. Junta de Andalucía. Consejería de Agricultura y Pesca. Dirección General de Desarrollo Forestal. Sevilla. 244pp.
- Thuiller W., Araújo M.B. & Lavorel S. (2003) Generalised models versus classification tree analysis: a comparative study for predicting spatial distributions of plant species at different spatial scales. *Journal of Vegetation Science*. 14: 669–680.
- Torres R.T., Santos J., Linnell J.D.C., Virgós E., Fonseca C. (2012)- Habitat use by sympatric red and roe deer in a Mediterranean ecosystem. *Animal Biology* 62: 351–366.

Vingada J., Fonseca C., Cancela J., Ferreira J. & Eira C. (2010). Ungulates and their management in Portugal. Pp.392–418. Em: Apollonio M., Andersen R. & Putman R. (eds). European Ungulates and their Management in the 21st Century. Cambridge University Press. 618pp.

Capítulo 2

Materiais e métodos

Demografia e modelação espacial

Capítulo 2

Materiais e métodos

A monitorização e o estudo da dinâmica populacional constituem um processo complexo. O conhecimento do tamanho, estrutura e composição de uma população permite avaliar como as variáveis intrínsecas (condição corporal, dieta, densidade, sanidade) e extrínsecas (coberto vegetal, clima, caça, predação, competição, entre outros) se auto-regulam e interagem. Coadjuvado a este aspeto surge a variabilidade espacial e como a distribuição de uma espécie/população é moldada pela expressão dos fatores ambientais que a influenciam.

Encontram-se descritos vários métodos de monitorização cuja adoção depende dos objetivos definidos e de uma série de fatores, designadamente tipo de *habitat*, orografia da área, ecologia e etologia da espécie, disponibilidade de recursos humanos e materiais, e nível de precisão e exatidão que pretendemos alcançar (Mayle *et al.* 1999). Os métodos são classificados em diretos, indiretos ou complementares (ver Mayle *et al.* 1999; McComb *et al.* 2010).

No caso particular dos cervídeos, são várias as metodologias aplicadas. Em *habitats* abertos privilegiam-se os métodos diretos (Granados *et al.* 2001; Acevedo *et al.* 2008), enquanto nas áreas de reduzida visibilidade são adotados métodos indiretos (e.g. contagem de excrementos (Putman 1984; Bugalho 1992; Ferreira 2002), quantificação de trilhos, cálculo de índices de consumo de vegetação (Acevedo *et al.* 2008), amostragem de distâncias para grupos de excrementos (Marques *et al.* 2001)), técnicas de captura-marcação-recaptura e radiotelemetria ou seguimento GPS.

No contexto do presente estudo, e dando seguimento aos trabalhos de monitorização realizados em anos transactos (Santos 2009; Carvalho 2011), adotou-se uma metodologia direta bastante difundida na monitorização de populações silvestres: definição de transectos lineares com aplicação da técnica de amostragem de distâncias. As metodologias directas, além de permitirem a obtenção de estimativas de densidades, fornecem-nos informação relativa à estrutura/composição populacional, rácio sexual e produtividade. Adicionalmente, a posição exata do grupo aquando da sua deteção será usada para discriminar quais os fatores ambientais que condicionam a presença/ausência da espécie e como estes afetam a sua organização social. Assim, a presente dissertação encontra-se dividida em duas abordagens distintas mas complementares: i) monitorização populacional e ii) análise de dados ambientais e interação espécie-*habitat*.

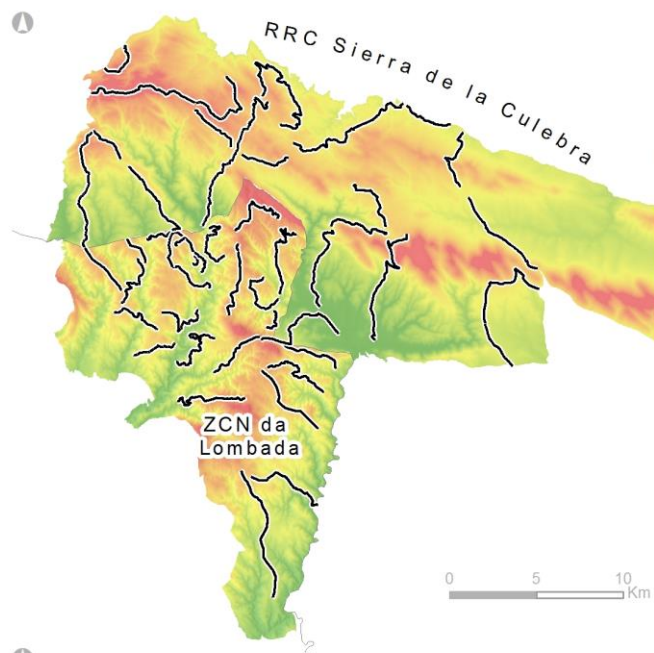
Demografia e monitorização populacional

Desenho experimental e periodicidade da amostragem

Procedeu-se à definição de uma rede de transectos para cada uma das áreas de estudo com o intuito de garantir a representatividade das mesmas e dos seus *habitats* característicos (Figura 5).

Os transectos, de conformação linear, foram percorridos por dois observadores registando-se o número de animais observados, o sexo, a classe etária e outras particularidades características de um indivíduo/grupo (Kie 1988). Para o cálculo posterior das densidades mediante o uso do *software Distance 6.0* (Thomas *et al.* 2009), marcaram-se os pontos GPS (*GarminTM Etrex Legend*) relativos à posição do observador e, com o auxílio de um telémetro (*LeicaTM Rangemaster CRF 1200*) e bússola, registaram-se as distâncias dos animais aos pontos de observação (r_i) e os respectivos azimutes: ângulo da direcção do transecto (θ_2) e o ângulo referente à posição do animal em relação ao observador (θ_1). A triangulação das medições permitiu-nos obter a posição exata do animal/grupo, bem como, a distância perpendicular ao transecto (x_i), essencial para a estimativa das densidades. Simultaneamente foi marcada a posição dos animais numa carta militar 1:10.000.

Os dados utilizados resultaram de diferentes amostragens nocturnas, devidamente autorizadas, realizadas com o auxílio de faróis de 100 Watts no período fenológico mais propício – a época de reprodução (brama), cujo pico é atingido nos meses de Setembro e Outubro, e no período pós-reprodutivo. A minimização do tempo despendido nas amostragens devido a aspectos fenológicos e etológicos da própria espécie (ver secção “Veado – breve caracterização da espécie” do capítulo 1) foi um aspeto tido em consideração.



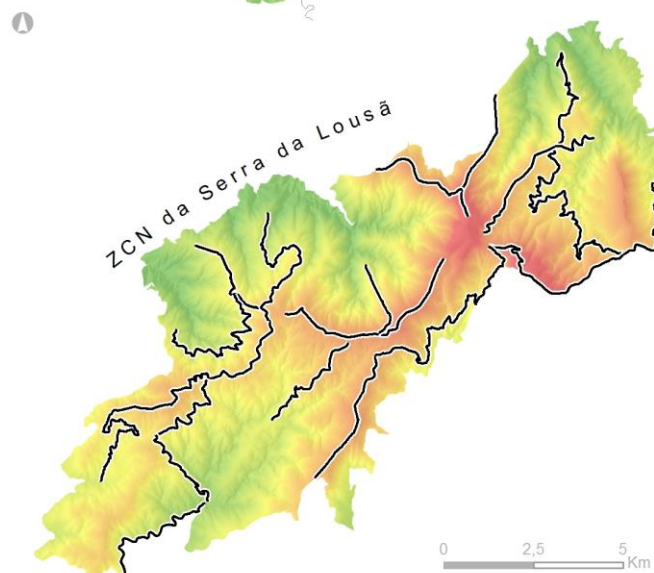
ZCN Lombada e RRC Sierra de la Culebra

Número total: 48 transectos
(28 ZCN Lombada e 20 RRC Sierra de la Culebra)

Extensão total: 223 Quilómetros
(103Km ZCN Lombada e
120Km RRC Sierra de la Culebra)

Extensão média: 4,6 Quilómetros
(3,7 (0,4 - 8,1) Km ZCN Lombada e
5,9 (2,4 - 12,7) Km RRC Sierra de la Culebra)

Número repetições: 2



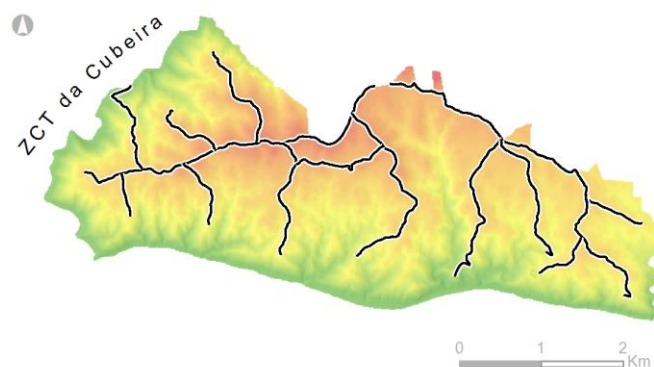
ZCN da Serra da Lousã

Número total: 20 transectos

Extensão total: 97 Quilómetros

Extensão média: 4,8 (1,7 - 9,6) Quilómetros

Número repetições: 2



ZCT da Cubeira

Número total: 14 transectos

Extensão total: 25 Quilómetros

Extensão média: 1,8 (0,5 - 4,2) Quilómetros

Número repetições: 2

Figura 5 – Disposição espacial dos transectos nas áreas de estudo.

Técnica de amostragem de distâncias – Distante 6.0

A técnica de amostragem de distâncias é bastante usada na estimativa de densidades e/ou abundâncias das populações de diversos *taxa* (Thomas *et al.* 2002) sendo aplicada, sobretudo, a percursos lineares (Thomas *et al.* 2010).

A robustez das estimativas encontra-se dependente do cumprimento de três premissas (Buckland *et al.* 1993; Thomas *et al.* 2010):

- i. Os animais que passem sobre a linha do transecto são sempre detectados ($g(0) = 1$);
- ii. Os animais são sempre detectados na sua posição inicial, antes de qualquer movimento em relação ao observador;
- iii. As distâncias perpendiculares (x_i) são medidas de forma exata.

Com recurso ao software *Distance 6.0* é obtida uma função de detecção $[f(x)]$ que nos fornece informação relativa à probabilidade de observar um grupo em função da distância perpendicular (x_i). Esta probabilidade geralmente decresce conforme aumenta a distância ao transecto (Kie 1988). A seleção da melhor função de detecção, obtida a partir da confrontação de funções-chave, termos de ajustamento e covariáveis com o histograma de frequências das distâncias perpendiculares, baseou-se no *Akaike's Information Criterion (AIC)* (Akaike 1974). A partir desta função foi estimada a densidade de animais presentes numa área de estudo (Buckland *et al.* 1993). Os dados das amostragens realizadas foram tratados recorrendo a dois motores de análise do programa, o *Conventional Distance Sampling (CDS)* e o *Multiple Covariates Distance Sampling (MCDS)*. No primeiro motor de análise, com o intuito de determinar o melhor modelo, testaram-se quatro funções-chave (*Uniform*, *Half-normal*, *Hazard-rate* e *Negative exponential*) com as séries de expansão disponíveis (*Cosine*, *Simple polynomial* e *Hermite polynomial*). O segundo motor de análise, MCDS, que apenas possui duas funções-chave (*Half-normal* e *Hazard-rate*) além das séries de expansão, permite a inclusão de covariáveis (Marques & Buckland 2003). Na análise foram efectuados truncamentos de distâncias e agruparam-se as distâncias perpendiculares em classes que permitissem a obtenção de uma função de detecção mais robusta.

CrITÉRIOS para avaliação da estrutura populacional

O registo dos dados populacionais foi realizado com base em classes de idade e sexo. Procedeu-se à definição das seguintes categorias:

Tabela 1 – Faixas etárias para avaliação da estrutura populacional. Baseado em Marco & Gortázar (2002).

Categorias (Faixa etária)	Intervalo de idades
Cria	Até 1 ano, independentemente do sexo
Fêmea jovem/subadulta (FJ)	1 a 3 anos
Fêmea adulta (FA)	A partir dos 4 anos
Vareto (MJ)	1 ano
Macho jovem/subadulto (MsubA)	2 a 4 anos
Macho adulto (MA)	5 a 9 anos
Macho velho (MV)	A partir dos 10 anos

As categorias de idade apresentadas são diferenciáveis no campo através do tamanho e forma do corpo, bem como, no caso dos machos, da configuração da armação.

Como critério para agrupamento dos dados utilizou-se o adotado por Carranza & Valencia (1992). Os grupos identificados completamente utilizaram-se em todos os cálculos. Os grupos identificados quanto ao género mas não à classe etária, foram utilizados para o cálculo do tamanho do tipo de grupo correspondente. Por último, os grupos nos quais não foi possível identificar o sexo nem idade dos constituintes, foram apenas utilizados na determinação dos tamanhos dos mesmos em geral.

Os grupos foram classificados de acordo com o descrito por Clutton-Brock *et al.* (1982) em grupos familiares (comumente constituídos por uma fêmea adulta, uma jovem/subadulta e/ou vareto e uma cria), grupos de machos ou grupos mistos.

Como critério para definir se um indivíduo pertence ou não a um grupo considerou-se uma distância mínima de 50 metros (Clutton-Brock *et al.* 1982). Porém, a análise comportamental dos animais permite, com maior certeza, associar ou não um animal a determinado grupo.

Um grupo consiste em um ou mais indivíduos e a determinação do seu tamanho constitui uma forma direta de caracterizar as unidades sociais (Reiczigel *et al.* 2008).

Krause & Ruxton (2002) demonstram que a generalidade dos grupos de aves e mamíferos são pequenos, sendo que o número de grupos grandes e extremamente grandes é reduzido. As distribuições enviesadas tornam a estatística descritiva inadequada para a análise dos dados (ver Rózsa *et al.* 2000). Jarman (1974) introduziu o conceito de Tamanho Típico do Grupo (do inglês *Typical Group Size*) que quantifica o tamanho médio do grupo experienciado individualmente. Assume o valor “1” para indivíduos solitários, o valor “2” para ambos os indivíduos de um grupo de dois e assim sucessivamente. Reiczigel *et al.* (2008) propõem uma nova metodologia aplicável a distribuições heterogêneas e não normais usando técnicas de *bootstrap* (*BCa bootstrap*) para o cálculo de intervalos de confiança. O conceito de aglomerado (do inglês *crowding*) é introduzido pelos autores como uma aproximação ao tamanho típico do grupo. O *software Flocker 1.1* foi usado no cálculo deste parâmetro (Reiczigel & Rózsa 2006). No presente estudo é utilizada a abordagem proposta por Reiczigel *et al.* (2008) para caracterizar as unidades sociais.

Análise de dados ecológicos e ambientais

A quantificação das relações espécie-*habitat* representa um dos maiores desafios da modelação ecológica e espacial (Guisan & Zimmermann 2000). Os dados usados em modelação resultam de amostragens sistemáticas com registos de presença/ausência (Hirzel & Guisan 2002) ou, na sua maioria, de registos unicamente de presença obtidos em coleções de herbário, museus ou resultantes de observações casuais (Huetmann 2005). Muitos autores têm ajustado os seus dados de presença a registos binários de presença/ausência mediante a criação de pseudo-ausências (Boyce *et al.* 2002). Atualmente, encontram-se descritas uma plethora de técnicas de modelação espacial com distintas especificidades e aplicáveis em diferentes contextos. A aplicação prática dos resultados de modelação está completamente dependente da discriminação precisa dos processos ecológicos e padrões espaciais que regem uma população. Segundo Austin (2007), a natureza complexa dos sistemas ecológicos diminui a nossa capacidade de gerar modelos precisos mediante a aplicação dos tradicionais métodos frequentistas. Assim, várias outras aproximações estatístico-matemáticas têm sido desenvolvidas.

As técnicas de modelação são comumente agrupadas em sete categorias distintas (Guisan & Zimmermann 2000):

- Regressões generalizadas e técnicas relacionadas – *e.g.* GLM, GAM;
- Técnicas de classificação – *e.g.* CART, MLC;
- *Environmental envelopes* – *e.g.* BIOCLIM, HABITAT, DOMAIN;
- Técnicas de ordenação – *e.g.* CCA, RDA;
- Aproximações *Bayesianas*;
- Redes neuronais – *e.g.* ANN;
- Outras aproximações – *e.g.* ENFA, GARP.

Quando são analisados padrões, a variação de uma variável resposta (*i.e.* dependente) é testada mediante a seleção de um conjunto de variáveis explicativas (*i.e.* independentes). A relação entre a primeira e as segundas é comumente medida através de coeficientes de regressão. A regressão múltipla é a técnica estatística mais usada para este propósito, contudo são várias as limitações das abordagens de regressão clássicas na análise de dados ecológicos (Carrascal *et al.* 2009). Os mesmos autores nomeiam três: i) os fenómenos ecológicos são geralmente descritos por um conjunto mais ou menos alargado de variáveis, as quais acresce o facto de, na maior

parte dos casos, serem redundantes, não serem independentes, exibirem interações complexas e autocorrelação; ii) o tamanho amostral poderá não ser suficientemente representativo quando comparado com o número de variáveis explicativas usadas, diminuindo a capacidade de uma regressão discriminar efeitos significativos, o que se traduz numa maior probabilidade de ocorrerem erros do tipo II; iii) os modelos de regressão clássicos não têm a capacidade de analisarem fenómenos complexos em que são consideradas, em conjunto, mais do que uma variável resposta. Em virtude destes constrangimentos, a regressão por mínimos quadrados parciais (*Partial Least Square Regression PLSR*) ou a própria análise multivariada de variância (*MANOVA*) constituem alternativas na análise de dados ecológicos.

As técnicas de *machine learning* (*ML*) têm ganho protagonismo na análise de dados ambientais/ecológicos. Ao contrário das abordagens estatísticas, estas técnicas assumem que o processo de geração de dados é complexo e incerto identificando a resposta dos mesmos mediante a discriminação de padrões dominantes (Elith *et al.* 2008). Nos últimos anos, as *CART* têm sido amplamente usadas devido à sua simples interpretação, à elevada precisão das análises que efetua e à capacidade de caracterizar interações complexas entre variáveis (Cutler *et al.* 2007). O seu poder analítico é ainda mais elevado mediante a aplicação de técnicas que criam e combinam os resultados de um conjunto numeroso de árvores de regressão (*e.g. Boosting Regression Trees* e *Random Forests*) (Evans *et al.* 2011).

Dados de presença/ausência

Os dados referentes às posições exactas dos animais/grupos identificados foram colectados ao longo de várias amostragens sistemáticas mediante o percurso de transectos lineares previamente definidos. A inexistência de dados precisos que registem a ausência da espécie pode ser contornada mediante a criação de pseudo-ausências (amostra de localizações sem informação relativa à presença/ausência da espécie) ou através de dados de *background* (amostra de localizações que representam a variação ambiental da área de estudo). No presente estudo usaram-se dados de *background* juntamente com os dados de presença no desenvolvimento e validação dos modelos (ver Peterson *et al.* 2011). O uso de dados de *background* requer poucos pressupostos no que respeita à sua seleção e tratamento estatístico (ver Phillips & Elith 2011). Para evitar que os modelos sejam sensíveis à prevalência da espécie, o rácio entre presenças/*background data* cifrou-se no 1:1 (Hirzel *et al.* 2006), contudo é frequente as

áreas de baixa adequabilidade serem mais representativas o que possibilita que o número de pseudo-ausências, quando estas são utilizadas, seja superior ao de presenças (Olivier & Wotherspoon 2006). Toda a informação encontra-se projetada no sistema global de referência recomendado pela *European Reference Frame (EUREF)* – o *European Terrestrial Reference System (ETRS89)*. A Tabela 2 sumariza o número de registos por área de estudo e época de amostragem.

Tabela 2 – Total de observações para as áreas de estudo por época de amostragem. EA – Época de acasalamento; EPA – Época pós-acasalamento.

Área de estudo	Época de amostragem	Período	Número de registos
ZCN da Lombada	EA 2010/2011	Dia	70
	EA 2010/2011	Noite	87
	EPA 2010/2011	Dia	64
	EA 2011/2012	Noite	64
	EA 2012/2013	Noite	78
RRC <i>Sierra de la Culebra</i>	EA 2011/2012	Noite	76
	EA 2012/2013	Noite	100
ZCN da Serra da Lousã	EA 2011/2012	Noite	67
	EA 2012/2013	Noite	85
	EPA 2012/2013	Noite	72
ZCT da Cubeira	EA 2010/2011	Dia	160
	EPA 2012/2013	Dia	65
TOTAL			988

Os modelos serão testados na população correspondente à ZCN da Lombada e RRC *Sierra de la Culebra*, por se considerar que esta constitui uma população modelo e porque os dados de base (variáveis) para esta área são fidedignos. É designada como população modelo devido ao tipo de gestão cinegética a que é sujeita e ao facto do seu principal predador (lobo-ibérico) permitir a manutenção de densidades populacionais equilibradas.

Variáveis ambientais

A seleção apropriada das variáveis ambientais é um factor crítico na *performance* preditiva dos modelos de adequabilidade ou seleção de *habitat*, contudo tem recebido menos atenção que a escolha das técnicas de modelação (Williams *et al.* 2012). O resultado da análise é igualmente influenciado pela resolução da informação e extensão da área de estudo (Anderson & Raza 2010). Idealmente, as variáveis deverão ser seleccionadas considerando as características ecológicas da espécie, a sua aplicação à escala da área de estudo e o seu poder preditivo (Post & Forchhammer 2002). Ao nível de pré-processamento, deverá proceder-se à sua reclassificação e ajustamento da resolução espacial.

No total foram seleccionadas 15 variáveis ambientais (Tabela 3) que, de acordo com bibliografia especializada (estrutura do *habitat* (Thomas *et al.* 1979; San-José *et al.* 1997; Debeljak *et al.* 2001), pressão humana (Hewinson *et al.* 2001) e disponibilidade hídrica (Thapaliya 2008)) e com o conhecimento empírico acerca da espécie, se encontram diretamente relacionadas com a distribuição do veado e o uso que a espécie faz do *habitat*. Atendendo ao número relativamente elevado de fatores ambientais relevantes para o estudo, o principal desafio passa pela definição de conjuntos de variáveis consistentes com os processos biofísicos (Franklin 2009).

Tabela 3 – Variáveis usadas no desenvolvimento dos modelos de classificação e regressão apresentados no capítulo 5.

Categoria	Variável	Resolução (m)
Disponibilidade hídrica	Distância a linhas de água (m)	30
Estrutura do habitat	% Campos agrícolas	
	% Floresta	
	% Matos	
Pressão humana	Distância a núcleos urbanos (m)	
	Distância a estradas (m)	
Produtividade	Índice de vegetação normalizado (NDVI)	
Topográficas	Declives (°)	
	<i>Northness</i>	
	<i>Eastness</i>	
	Altimetria (m)	
	Rugosidade do terreno	
Uso do solo	Distância a florestas (m)	
	Distância a campos agrícolas (m)	
	Distância a áreas de matos (m)	

A exposição é uma variável circular, pelo que deverá ser transformada recorrendo a funções trigonométricas seno e co-seno (Zar 1999). O resultado são duas novas variáveis que quantificam o grau de exposição de uma vertente a norte (*Northness*) e a este (*Eastness*).

$$Northness = \text{seno} \left(\frac{\text{exposição} * \pi}{180} \right)$$

$$Eastness = \text{coseno} \left(\frac{\text{exposicao} * \pi}{180} \right)$$

O índice de vegetação normalizado (*NDVI*) foi obtido com recurso às bandas 3 (*Red*) e 4 (*NIR*) da imagem de satélite *Landsat 7 ETM+*.

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

A exclusão de variáveis relativas à predação e competição intra e/ou interespecífica deveu-se à inexistência de dados.

Análise exploratória e técnicas de modelação

A análise exploratória dos dados é um procedimento essencial e que vai determinar a seleção do método de modelação a aplicar. Constitui o primeiro passo na análise de dados (Zuur *et al.* 2007). Apesar de algumas violações aos pressupostos estatísticos terem pouco impacto, a análise exploratória permitirá a redução da probabilidade de erro. A análise foi efetuada seguindo os procedimentos constantes em Zuur *et al.* (2010). Os métodos aplicados são robustos e pouco sensíveis à distribuição dos dados, contudo, a exploração dos mesmos possibilitará a análise de correlações entre variáveis e a prevenção/justificação de eventuais enviesamentos e reduzida *performance* dos modelos desenvolvidos. É igualmente útil na análise de dados complexos como os ambientais que se caracterizam pela sua elevada dimensionalidade, relações não-lineares, correlação espacial e heterocedasticidade.

Estudos recentes demonstram que as projeções de vários modelos para descrever um mesmo fenómeno são distintas (ver Thuiller *et al.* 2004; Araújo *et al.* 2005). Assim, seguindo a abordagem proposta por Araújo & New (2007), pretende-se aplicar um conjunto de técnicas de modelação que se traduzam em informação mais consistente.

Proceder-se-á a uma descrição sucinta dos métodos de modelação aplicados no contexto da dissertação, sendo que, a descrição mais específica é efetuada nos capítulos correspondentes.

A análise multicritério ponderada pelo método analítico hierárquico (capítulo 4) é um método baseado no conceito de “*expert knowledge*” de grande utilidade nos estudos de aptidão ambiental uma vez que constitui um procedimento racional que permite estruturar um problema constituído por vários elementos de natureza distinta. A

quantificação dos diferentes elementos e a conversão do conjunto de informação em valores numéricos permite o seu processamento e comparação sobre toda a sua extensão. O principal objectivo do método corresponde à escolha da melhor solução que satisfaça os valores de entrada (fatores) e as restrições. Medido através de uma escala de valores, os fatores poderão potenciar ou minimizar a aptidão de uma alternativa para o resultado final. A atribuição de pesos é realizada seguindo uma escala numérica previamente estabelecida (Saaty & Vargas 1991). Mais informações acerca da atribuição de ponderações, aplicação e validação do modelo são fornecidas no capítulo 4.

No capítulo 5 é testada a aplicabilidade de quatro metodologias (*CART*, *RF*, *BRT* e *GLM*), avaliada a sua capacidade preditiva e explicativa e testada a sua transferibilidade espacial.

As *CART* são um método flexível que modelam eficazmente interações complexas entre variáveis e cujo resultado final é facilmente interpretável (Breiman *et al.* 1984; De'Ath & Fabricius 2000). O desenvolvimento de um modelo *CART* é realizado em quatro passos principais: i) partição dos dados-treino; ii) ajuste do modelo considerando partições prévias dos dados; iii) interrupção do algoritmo quando os resíduos do modelo são aproximadamente zero e o número de variáveis que restam é reduzida; iv) simplificação da árvore de regressão através do parâmetro de complexidade para evitar um sobreajuste aos dados.

O conceito de *ensemble* é introduzido pelos algoritmos *RF* e *BRT*. *RF* é um método onde cada árvore é construída independentemente através de um conjunto de amostras de dados obtidas por técnicas *bootstrap* (é considerado um *bagging-based method*) (Breiman 1996). As novas árvores não dependem das anteriores. O modelo final é obtido considerando a média dos resultados dos submodelos, as árvores de classificação ou regressão individuais. Este método retém as variáveis que providenciam mais informação na discriminação de padrões (Evans *et al.* 2011).

As *BRT* consistem numa sequência de modelos cuja árvore posterior altera o peso das observações prévias, no sentido de dar mais relevância aos registos classificados incorrectamente (De'Ath 2007). O objetivo das *BRT* é melhorar a *performance* do modelo através da combinação de um número elevado de árvores evitando a definição de uma única predição (Elith *et al.* 2008).

O *GLM* (McCullagh & Nelder 1989) constitui uma generalização flexível da regressão linear no que concerne ao tipo de distribuição da variável resposta, permitindo várias distribuições além da clássica distribuição normal. Estes modelos de ampla aplicabilidade providenciam uma descrição das relações entre uma variável resposta, dicotômica ou não, e um conjunto de variáveis explicativas.

Reclassificação e representação cartográfica

No capítulo 4, o modelo final obtido foi reclassificado em classes de adequabilidade através da aplicação do método de intervalos naturais, um dos mais comuns na classificação de dados quantitativos (Conolly & Lake 2006). Na presença de dados contínuos, este método agrega os mesmos em conjuntos/classes lógicos mediante a análise de toda a distribuição e identificação de pontos de quebra cuja finalidade é aumentar a diferença entre os números de classes definidos (Jenks & Caspall 1971). Segundo Hirzel *et al.* (2006), a definição de um número reduzido de classes facilita informação mais robusta, pelo que se procedeu à reclassificação do modelo em apenas três classes de adequabilidade. No capítulo 5, a reclassificação foi efetuada manualmente definindo-se intervalos iguais de probabilidade o que facilita a comparação cartográfica e visual entre os diferentes métodos testados.

Avaliação da *performance* dos modelos

A validação dos modelos pode ser realizada segundo duas abordagens distintas. A primeira mediante o uso dos mesmos dados com os quais se construiu o modelo, através da aplicação de técnicas *bootstrap*, validação cruzada (*cross-validation*) ou o método de *Jackknife*. A segunda, através de dados independentes e comparação de valores usando a curva *ROC* (*Receiver Operating Characteristic*) e gráficos de calibração.

A curva *ROC* permite testar a precisão do modelo proporcionando uma validação independente de distorções e potenciais enviesamentos (Fielding & Bell 1997). O valor obtido varia entre 0,50 e 1, sendo que o primeiro expressa a sobreposição dos valores dos dois grupos enquanto o segundo indica que estes não coincidem. A capacidade de discriminação dos modelos (capacidade do modelo em diferenciar os registos de presença dos de ausência) é, usualmente, a única característica usada na avaliação dos modelos preditivos (Jiménez-Valverde *et al.* 2013), contudo, a sua calibração (concordância entre as probabilidades de ocorrência preditas e a proporção observada de

presenças) é outro aspecto que providencia importante informação acerca da *performance* preditiva dos modelos (ver Pearce & Ferrier 2000; Jiménez-Valverde *et al.* 2013). No capítulo 4, artigo já publicado, a validação é efetuada com base na curva *ROC*, enquanto no capítulo 5, artigo em preparação, procede-se a uma avaliação preditiva (curva *ROC* e gráficos de calibração) e explicativa (correlação entre o grau de importância das variáveis entre os vários modelos desenvolvidos) dos modelos.

A sequência dos principais procedimentos adotados é esquematicamente apresentada na Figura 6.

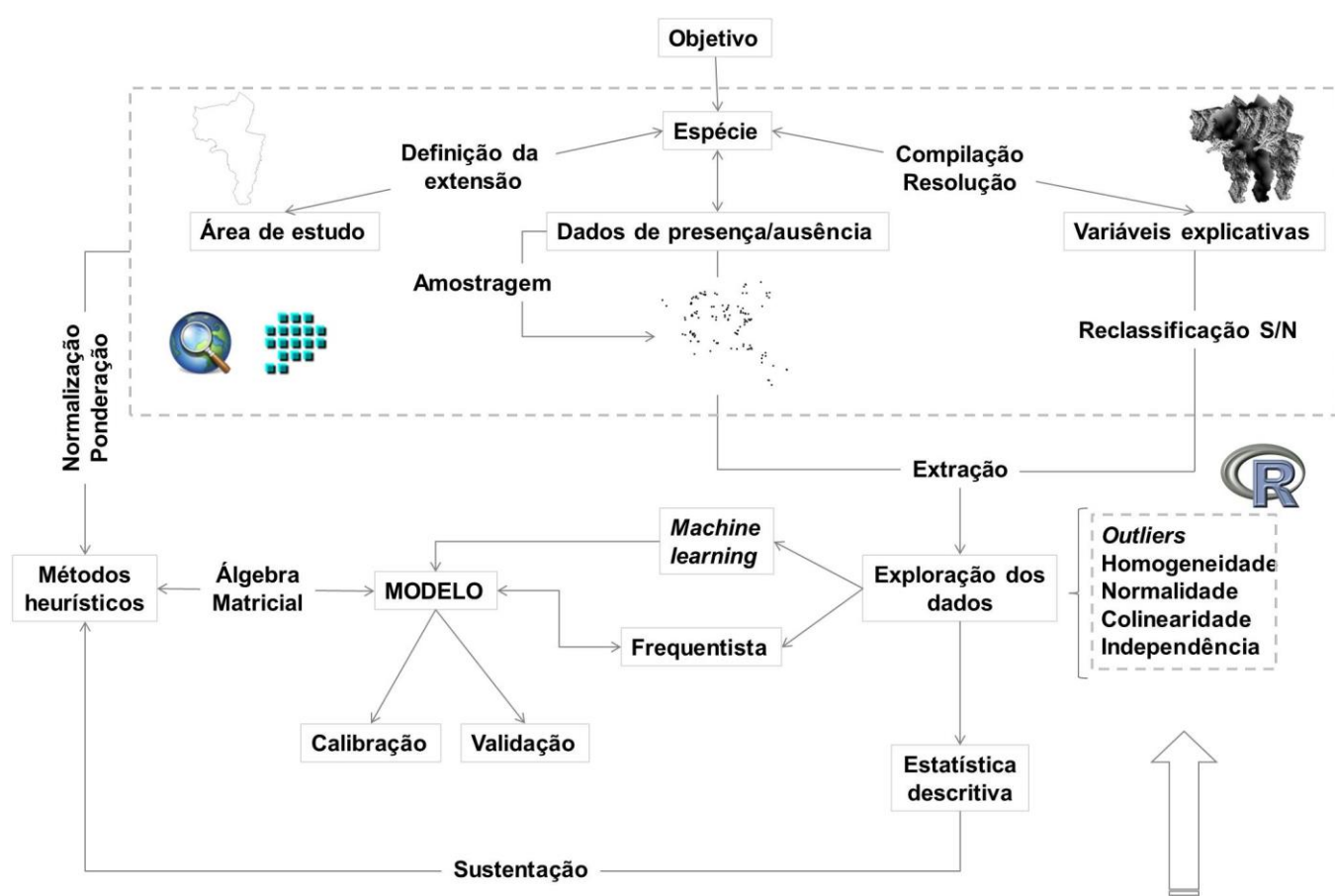


Figura 6 – Sequência dos principais passos adotados no desenvolvimento dos modelos apresentados no contexto da dissertação.

Transferibilidade no espaço

A transferibilidade espacial de modelos pode ser problemática uma vez que se torna difícil avaliar a *performance* preditiva dos mesmos para as áreas em que estes são projetados (Barbosa *et al.* 2009). Se os modelos são transferidos entre áreas com características ambientais similares e se a espécie estudada não exibe uma plasticidade ambiental significativa, os resultados são robustos (Bulluck *et al.* 2006). Os principais problemas dizem respeito à variabilidade ambiental entre as áreas de estudo e à qualidade dos dados-treino e teste. Mesmo que a homogeneidade das condições abióticas seja garantida, a distinta influência histórica na distribuição da espécie pode condicionar a transferibilidade dos modelos desenvolvidos (Randin *et al.* 2006).

No presente estudo (capítulo 5) é testada a transferibilidade espacial dos quatro modelos desenvolvidos (*CART*, *RF*, *BRT* e *GLM*) e avaliada a sua *performance* com recurso a dados de presença da espécie para a área de transferência. O processo foi executado em quatro passos: i) aplicação dos *Variance Inflation Factors* (VIF) para quantificar a multicolinearidade das variáveis; ii) avaliação da similaridade ambiental entre as áreas de estudo com recurso à distância de *Mahalanobis* (ver Cianfrani *et al. in press*); iii) avaliação da correlação entre as duas matrizes ambientais através da significância estatística do teste de *Mantel*; iv) avaliação da *performance* preditiva dos modelos quando aplicados na área de transferência.

Referências bibliográficas

- Acevedo P., Ruiz-Fons F., Vicente J., Reyes-García A.R., Alzaga V. & Gortázar C. (2008). Estimating red deer abundance in a wide range of management situations in Mediterranean habitats. *Journal of Zoology*. 276: 37–47.
- Akaike H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 19: 716–723.
- Anderson R.P. & Raza A. (2010). The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*. 37: 1378–1393.
- Araújo M.B., Whittaker R.J., Ladle R.J. & Erhard M. (2005). Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*. 14: 529–538.
- Araújo M.B. & New M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*. 22: 42–47.
- Austin M. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*. 200: 1–19.
- Barbosa A.M., Real R. & Vargas J.M. (2008). Transferability of environmental favourability models in geographic space: The case of the Iberian desman (*Galemys pyrenaicus*) in Portugal and Spain. *Ecological Modelling*. 220: 747–754.
- Boyce M.S., Vernier P.R., Nielsen S.E. & Schmiegelow F.K.A. (2002). Evaluating resource selection functions. *Ecological Modelling*. 157: 281–300.
- Breiman L., Friedman J.H., Olshen R.A. & Stone C.J. (1984). Classification and regression trees. Wadsworth International Group. 358pp.
- Breiman L. (1996). Bagging predictors. *Machine Learning*. 24(2): 123–140.

Buckland S.T., Ahmadi S., Staines W., Gordon I.J. & Youngson R.W. (1993). Estimating the minimum population size that allows a given annual number of mature red deer stags to be culled sustainability. *Journal of Applied Ecology*. 33: 118–190.

Bugalho M.N. (1992). Aplicação do método da contagem de grupos de excrementos para determinação do uso de habitat e densidades populacionais numa população de veados (*Cervus elaphus* L.). Relatório do trabalho de fim do curso de Engenharia Florestal. Instituto Superior de Agronomia. Universidade Técnica de Lisboa.

Bulluck L., Fleishman E., Betrus C. & Blair R. (2006). Spatial and temporal variations in species occurrence rate affect the accuracy of occurrence models. *Global Ecology and Biogeography*. 15: 27–38.

Carranza J. & Valencia J. (1992). Organización social del ciervo en hábitat mediterráneo. *Miscellanea Zoologica*. 16: 223–232.

Carrascal L.M., Galván I. & Gordo O. (2009). Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos*. 118: 681–690.

Carvalho J. (2011). Monitorização da população de veado (*Cervus elaphus* L.) na Zona de Caça Nacional da Lombada, Nordeste Transmontano: Época 2010/2011. Relatório de Estágio para obtenção da licenciatura em Biologia. Universidade de Aveiro.

Cianfrani C., Maiorano L., Loy A., Kranz A., Lehmann A., Maggini R. & Guisan A. (*in press*). There and back again? Combining habitat suitability modelling and connectivity analyses to assess a potential return of the otter to Switzerland. *Animal Conservation*.

Clutton-Brock T.H., Guinness F.E. & Albon S.D. (1982). Red deer: Behavior and Ecology of two sexes. Edinburgh University Press. Edinburgh. 400pp.

Conolly J. & Lake M. (2006). Geographical Information Systems in Archaeology. Cambridge Manuals in Archaeology. Cambridge University Press. United Kingdom. 358 pp.

Cutler D.R., Edwards T.C., Beard K.H., Cutler A., Hess K.T., Gibson J. & Lawler J.J. (2007). Random forests for classification in ecology. *Ecology*. 88(11): 2783–2792.

De'ath G. & Fabricius K.E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*. 81(11): 3178–3192.

De'ath G. (2007). Boosted trees for ecological modelling and prediction. *Ecology*. 88(1): 243–251.

Debeljak M., Saso D., Klemen J., Andrej K. & Adamie M. (2001). Habitat suitability modelling for red deer (*Cervus elaphus* L.) in South-central Slovenia with classification trees. *Ecological Modelling*. 138(1-3): 321–330.

Elith J., Leathwick J.R. & Hastie T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*. 77: 802–813.

Evans S.E., Murphy M.A., Holden Z.A. & Cushman S.A. (2011). Modelling species distribution and change using Random Forest. Pp.139-160. Em: Drew C.A., Wiersma Y.F. & Huettmann F. (eds.). *Predictive Species and Habitat Modelling in Landscape Ecology – Concepts and Applications*. Springer. 313pp.

Ferreira J.A. (2002). Caracterização da população de corço (*Capreolus capreolus*) no Parque Nacional da Peneda-Gerês. Instituto da Conservação da Natureza. 37pp.

Fielding A.H. & Bell J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*. 24: 38–49.

Franklin J. (2009). Mapping species distributions: spatial inference and prediction. Cambridge. Cambridge University Press. 338pp.

Granados J.E., Cabrera M.C., García J. & Fandos P. (2001). El Ciervo común (*Cervus elaphus* Linnaeus, 1758) en el Parque Natural de Sierra de Baza (Granada). *Galemys*. 13: 27–37.

Guisan A. & Zimmermann N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*. 135: 147–186.

Hewinson A.J.M., Vincent J.P., Joachim J., Angibault I.M., Cargnelutti B. & Cibiën C. (2001). The effects of woodland fragmentation and human activity on roe deer distribution in agricultural landscapes. *Canadian Journal of Zoology / Revue Canadienne de Zoologie*, 79(4): 679–689.

Hirzel A.H. & Guisan A. (2002). Which is the optimal sampling strategy for habitat suitability modelling? *Ecological Modelling*. 157: 331–341.

Hirzel A.H., Le Lay G., Helfer V., Randin C. & Guisan A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*. 199: 142–152.

Huettmann F. (2005). Databases and science-based management in the context of wildlife and habitat: towards a certified ISO standard for objective decision-making for the global community by using the internet. *Journal of Wildlife Management*. 69: 466–472.

Jarman P.J. (1974). The social organisation of antelope in relation to their ecology. *Behaviour*. 48: 215–267.

Jenks G.F. & Caspall F.C. (1971). Error on choroplethic maps: definition, measurement, reduction. *Annals of the Association of American Geographers*. 61: 217–244.

Jiménez-Valverde A., Acevedo P., Márcia Barbosa A., Lobo J.M. & Real R. (2013). Discrimination capacity in species distribution models depends on the representativeness of the environmental domain. *Global Ecology and Biogeography*. 22: 508–516.

Kie J.G. (1988). Performance in wild ungulates: measuring population density and condition of individuals. Gen. Tech. Rep. PSW – 106. Berkeley, CA: Pacific Southwest Forest and Range Experiment Station. Forest Service. United States Department of Agriculture. 17pp.

Krause J. & Ruxton G.D. (2002). *Living in groups*. Oxford University Press. 228pp.

Marco J. & Gortázar C. (2002). El ciervo: Biología y gestión poblacional. Prames, Zaragoza. 35pp.

Marques F.F.C., Buckland S.T., Goffin D., Dixon C.E., Borchers D.L., Mayle B.A & Peace A.J. (2001). Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland. *Journal of Applied Ecology*. 38: 349–363.

Marques F.F.C. & Buckland S.T. (2003). Incorporating covariates into standard line transect analyses. *Biometrics*. 59: 924–935.

Mayle B.A., Peace A.J. e Gill R.M. (1999). How many deer? A field guide to estimating deer population size. Edinburgh: Forestry Commission Field Book 18. Forestry Commission. 96pp.

McComb B., Zuckerber B., Vesely D. & Jordan C. (2010). Monitoring Animal Populations and their Habitat. A Practitioner's Guide. CRC Press. 299pp.

McCullagh P. & Nelder J.A. (1989). Generalized Linear Models. London: Chapman and Hall. 511pp.

Olivier F. & Wotherspoon S.J. (2006). Modelling habitat selection using presence-only data: Case study of a colonial hollow nesting bird, the snow petrel. *Ecological Modelling*. 195: 187–204.

Pearce J. & Ferrier S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*. 133: 225–245.

Peterson A.T., Soberón J., Pearson R.G., Anderson R.P., Martínez-Mayer E., Nakamura M. & Araújo M.B. (2011). Ecological niches and geographic distribution. Princeton University Press. 328pp.

Phillips S.J. & Elith J. (2011). Logistic methods for resource selection functions and presence-only species distribution models, AAAI (Association for the Advancement of Artificial Intelligence). San Francisco, USA. 1384–1389.

Post E. & Forchhammer M.C. (2002). Synchronization of animal population dynamics by large-scale climate. *Nature*. 420: 168–171.

Putman R.J. (1984). Facts from faeces. *Mammal Review*. 14: 79–97.

Randin C.F., Dirnböck T., Dullinger S., Zimmermann N.E., Zappa M. & Guisan A. (2006). Are niche-based species distribution models transferable in space? *Journal of Biogeography*. 33: 1689–1703.

Reiczigel J. & Rózsa L. (2006). Flocker 1.1. Disponível em: <http://www.zoologia.hu/flocker/flocker.html>

Reiczigel J., Lang Z., Rózsa L. & Tóthmérész B. (2008). Measures of sociality: two different views of group size. *Animal Behaviour*. 75: 715–721.

Rózsa L., Reiczigel J. & Majoros G. (2000). Quantifying parasites in samples of hosts. *Journal of Parasitology*. 86: 228–232.

Saaty T.L. & Vargas L.G. (1991). *Prediction, Projection and Forecasting*. Kluwer Academic Publisher. 251 pp.

San-José C., Braza F., Aragón S. & Delibes J.R. (1997). Habitat use by roe and red deer in southern Spain. *Miscellania Zoologica*. 20: 27–38.

Santos J. (2009). Estudo populacional do veado (*Cervus elaphus* L.) no Nordeste Transmontano. Tese de Mestrado em Ecologia, Biodiversidade e Gestão de Ecossistemas. Universidade de Aveiro.

Thapaliya K. (2008). Analysis of factors related to the distribution of red deer (*Cervus elaphus* L.) in Hustai National Park, Mongolia. M.Sc. Thesis. International Institute for Geo-Information Science and Earth Observation. 50pp.

Thomas J.W., Black H., Scherzinger R.J. & Pedersen R.J. (1979). Deer and Elk. Pp. 104-127. In: J.W. Thomas (ed). Wildlife habitats in managed forests of the Blue Mountains of Oregon and Washington, Agricultural Handbook. Wildlife Management Institute, Washington, DC, USA.

Thomas L., Buckland S.T., Burnham K.P., Anderson D.R., Laake J.L., Borchers D.L. & Strindberg S. (2002). Pp.544-552. Distance Sampling. Em: El-Shaarawi, A.H & Piegorisch W.W. (eds.). Encyclopedia of Environmetrics. John Wiley & Sons. Volume 1.

Thomas L., Laake J.L., Rexstad E., Strindberg, S., Marques, F.F.C., Buckland S.T., Borchers D.L., Anderson D.R., Burnham K.P., Burt M.L., Hedley S.L., Pollard J.H., Bishop J.R.B. and Marques T.A. (2009). Distance 6.0. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. <http://www.ruwpa.st-and.ac.uk/distance/>

Thomas L., Buckland S.T., Rexland E.A., Laake J.L., Strindberg S., Hedley S.L., Bishop J.R.B., Marques T.A. & Burnham K.P. (2010). Distance software: design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology. 47: 5–14.

Thuiller W., Araújo M.B., Pearson R.G., Whittaker R.J., Brotons L. & Lavorel S. (2004). Biodiversity conservation: Uncertainty in predictions of extinction risk. Nature. 430: 10.1038/nature02716.

Williams K.J., Belbin L., Austin M.P., Stein J.L. & Ferrier S. (2012). Which environmental variables should I use in my biodiversity model? International Journal of Geographical Information Science. 26(11): 2009–2047.

Zar J.H. (1999). Biostatistical Analysis. 5th edition. Prentice-Hall. 859pp.

Zuur A.F., Ieno E.N. & Smith G.M. (2007). Analysing Ecological Data. Springer. 672pp.

Zuur A.F., Ieno E.N. & Elphick C.S. (2010). A protocol for data exploration to avoid common statistical problems. Methods in Ecology & Evolution. 1: 3–14.

Capítulo 3

Resultados e Discussão

Demografia e Monitorização populacional

Capítulo 3

Caraterização demográfica

Estrutura populacional e parâmetros demográficos

No que concerne à estrutura populacional, para a população do Noroeste da Península Ibérica (ZCN Lombada e RRC *Sierra de la Culebra*) obtiveram-se os dados apresentados na Figura 7. Atendendo ao facto de as duas zonas de caça estarem sujeitas a regimes de gestão cinegética distintos, o que se pode refletir na estrutura populacional, os dados referentes a estas são apresentados em separado.

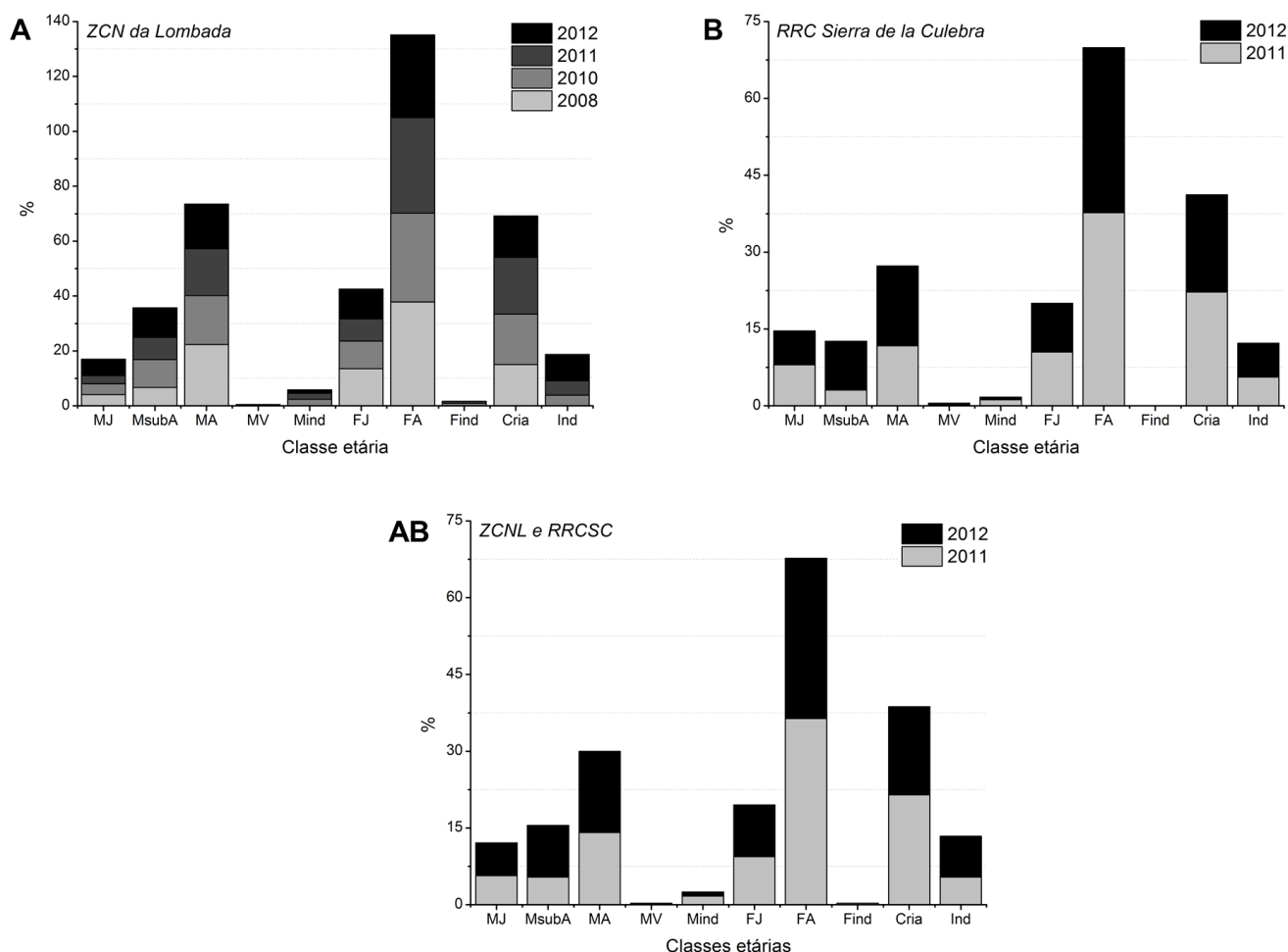


Figura 7 – Estrutura da população de veado do Noroeste da Península Ibérica. A) Dados relativos a quatro anos de amostragem na ZCN da Lombada; B) Dados relativos a dois anos de amostragem na RRC *Sierra de la Culebra*; AB) Agrupamento dos dados das duas populações. (MJ) – Vareto; (MsubA) – Macho jovem/subadulto; (MA) – Macho adulto; (MV) – Macho velho; (Mind) – Macho indeterminado; (FJ) – Fêmea jovem; (FA) – Fêmea adulta; (Find) – Fêmea indeterminada; (Ind) – Indeterminado.

Os dados das monitorizações demonstram um equilíbrio ao longo das várias amostragens e um equilíbrio da própria estrutura populacional. O predomínio de exemplares adultos, mais evidente na ZCN da Lombada, é elucidativo da reduzida atividade cinegética a que a população está sujeita, apesar do furtivismo que se regista. As fêmeas adultas mantêm-se como a classe etária mais representativa, o que é perfeitamente natural mesmo em populações não sujeitas a pressão cinegética, face à sua maior longevidade comparativamente com os machos da espécie e ao facto desta classe etária ser representada por animais com idade superior a quatro anos, tornando-a a mais abrangente. O número de machos velhos observados é um reflexo do seu reduzido número e/ou adoção de comportamentos que dificultam a sua observação (Santos 2009). Clutton-Brock *et al.* (1982) refere que a mortalidade de crias do sexo masculino é maior desde o momento da sua concepção até ao primeiro ano de vida o que, se reflete no rácio sexual de uma população. Outra questão importante é a sazonalidade demarcada dos padrões comportamentais da espécie que pode contribuir para as diferenças observadas no que concerne a este rácio em particular. Os machos, após a época de acasalamento, tendem a dispersar e a refugiarem-se em áreas que proporcionem alimento suficientemente rico (e.g. bolota e castanha) que lhes permita recuperar o peso corporal entretanto perdido (Barroso & Rosa 1999), evitando, igualmente, entrar em competição com as fêmeas pelas melhores áreas de alimentação. Destes *habitats* destacam-se os carvalhais e os soutos que, devido às suas características, dificultam a observação dos animais. Apesar dos dados apresentados dizerem respeito ao período de acasalamento, a extensão do período de amostragem e sua sobreposição a períodos fenológicos distintos, pode repercutir-se num ligeiro enviesamento dos resultados obtidos.

As populações da serra da Lousã e da ZCT da Cubeira caracterizam-se por discrepâncias mais vincadas no que concerne à estrutura populacional (Figura 8).

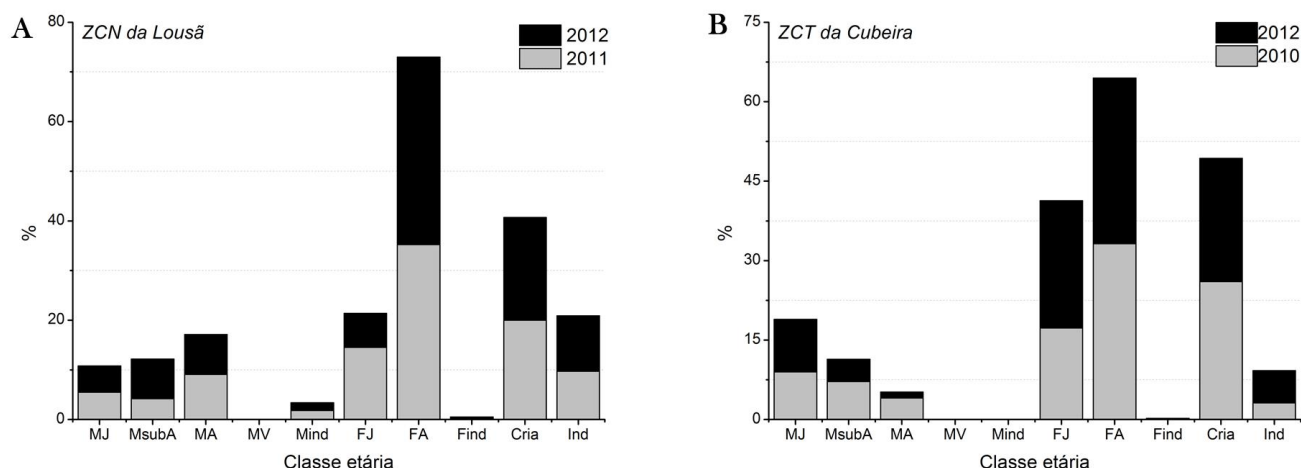


Figura 8 – Estrutura da população de veado para a ZCN da Lousã (A) e para a ZCT da Cubeira (B). (MJ) – Vareto; (MsubA) – Macho jovem/subadulto; (MA) – Macho adulto; (MV) – Macho velho; (Mind) – Macho indeterminado; (FJ) – Fêmea jovem; (FA) – Fêmea adulta; (Find) – Fêmea indeterminada; (Ind) – Indeterminado.

Os modelos de gestão adotados e os objetivos da exploração cinegética praticada são parte da justificação destes valores que revelam um evidente desequilíbrio. Além destes fatores, realçam-se a periodicidade da amostragem que, por motivos logísticos, não pôde ser realizada no período correspondente ao pico da brama em ambas as populações. Ecologicamente, algumas razões poderão estar na base do enviesamento favorável às fêmeas. Muitas populações animais apresentam enviesamentos no rácio sexual à nascença, contudo as causas são multivariadas sendo que, a preponderância das mesmas, depende das condições ambientais (Kruuk *et al.* 1999). Os mesmos autores demonstram que a diminuição do número de machos à nascença é um reflexo do aumento da densidade populacional e da precipitação total invernal. Estes dois aspetos encontram-se relacionados com o *stress* nutricional durante o período de gestação. Trivers & Willard (1973) documentam que as fêmeas com uma condição física deficitária tendem a produzir um rácio reduzido entre crias macho e fêmea. Na ilha de Rum, Escócia, ficou demonstrado que as fêmeas dominantes produziam, consistentemente, mais crias macho do que fêmeas, sendo que, o aumento da densidade populacional levou a um declínio da proporção de machos nascidos anualmente (Clutton-Brock *et al.* 1984).

A conjugação dos modelos de gestão, periodicidade de amostragem e dos aspetos biológicos da própria espécie, constituem as razões chave para o desequilíbrio

da estrutura etária observada para as três populações. Será, contudo, necessário realizar amostragens no pico do período de acasalamento da espécie para as três populações no sentido de aferir e justificar, com maior certeza, as diferenças observadas.

A caracterização demográfica de uma população envolve a determinação de vários indicadores que são elucidativos do equilíbrio, estado e produtividade da população. Na Tabela 4 são apresentados dados relativos à densidade, rácio sexual e produtividade para as três populações.

Tabela 4 – Dados relativos à densidade, rácio sexual (proporção macho/fêmea) e produtividade (proporção cria/fêmea) para as três populações. (*) Valores obtidos por Santos (2009). (**) Valores observados que poderão não corresponder ao rácio anual efetivo devido ao período em que foi realizada a amostragem. As variações no valor de densidade foram calculadas para um intervalo de confiança de 95%.

	Densidade (veados/100ha)	Rácio sexual (rs)	Produtividade (pr)
ZCN da Lombada	3,26 (2,27 – 4,70)*	0,78 (1:1,28)	0,36 (1:2,77)
ZCN da Lousã	5,52 (2,90 – 10,4)	0,32 (1:3,08)**	0,40 (1:2,48)
ZCT da Cubeira	23,49 (15,83 – 34,85)	0,28 (1:3,63)**	0,42 (1:2,23)

No que concerne à densidade populacional serão discutidos os valores obtidos no âmbito do presente estudo referentes às populações da ZCN da Lousã e ZCT da Cubeira. Para a ZCN da Lombada, a discussão dos valores de densidade estimados é devidamente aprofundada em Santos (2009) e Carvalho (2011).

As estimativas de densidade foram realizadas com recurso ao *software Distance 6.0* (Thomas *et al.* 2009). Como base para a seleção do melhor modelo foi tido em consideração o menor valor de *AIC* (*Akaike's Information Criterion*) (Akaike 1974), complementado, sempre que justificável, com os valores de ΔAIC e χ^2 *Goodness-of-fit* (teste de ajustamento Qui-quadrado). Para a ZCN da Lousã o modelo mais ajustado foi o *Negative Exponential* com série de expansão *Cosine* sem qualquer covariável associada (*AIC*=193,36 e coeficiente de variação de 33,15%). As distâncias perpendiculares foram agrupadas em intervalos de 20 metros procedendo-se a dois truncamentos, um à direita aos 140 metros e outro à esquerda, descartando-se as observações cuja distância perpendicular era inferior a 20 metros (Figura 9).

A densidade média estimada para a ZCN da Lousã no período invernal cifrou-se nos 5,52 veados/100ha. O coeficiente de variação obtido contribui para alguma imprecisão da estimativa pois, para um intervalo de confiança de 95%, o intervalo de variação da densidade está compreendido entre os 2,91 e os 10,50 veados/100ha. Os valores obtidos traduziram-se num total de aproximadamente 610 (317 - 1144) animais para a extensão da área de estudo no período invernal.

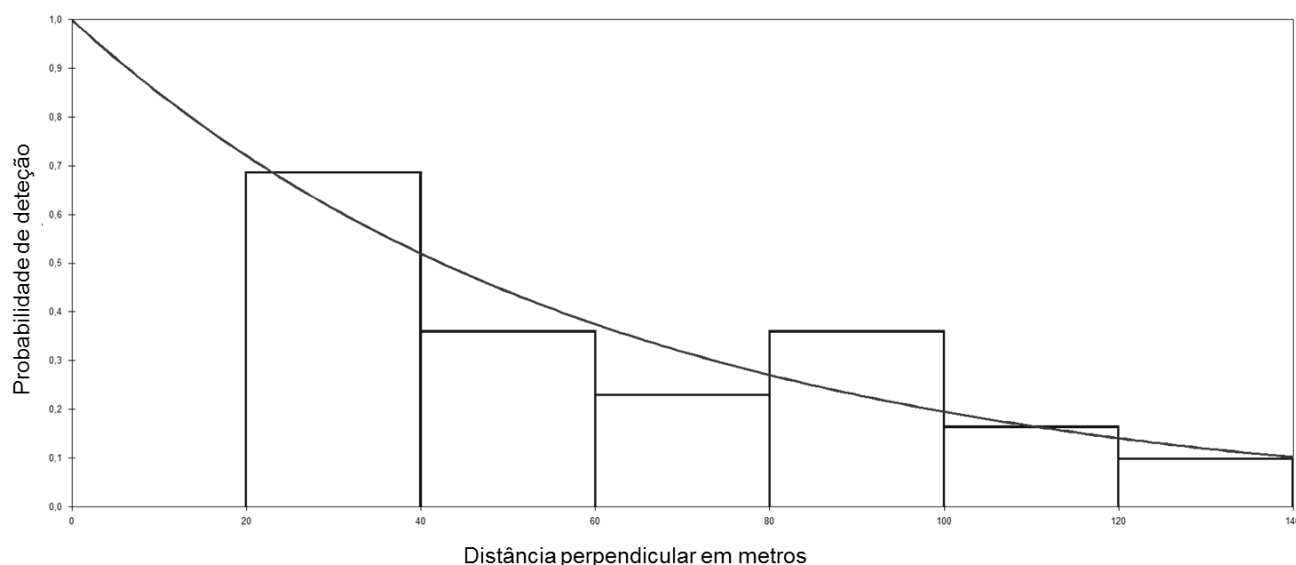


Figura 9 – Histograma que constitui a representação gráfica da função de detecção para a ZCN da Lousã ilustrando a probabilidade de detectar veados em função da distância perpendicular.

Relativamente à ZCT da Cubeira, o modelo que mais se ajustou aos dados foi desenvolvido com a função-chave *Half-Normal* associada à série de expansão *Cosine* sem covariáveis seleccionadas ($AIC=126,80$ e coeficiente de variação de 19,7%). Procedeu-se ao agrupamento das distâncias perpendiculares em intervalos de 20 metros e ao truncamento à direita aos 120 metros garantido um bom ajuste da curva teórica aos dados registados (Figura 10).

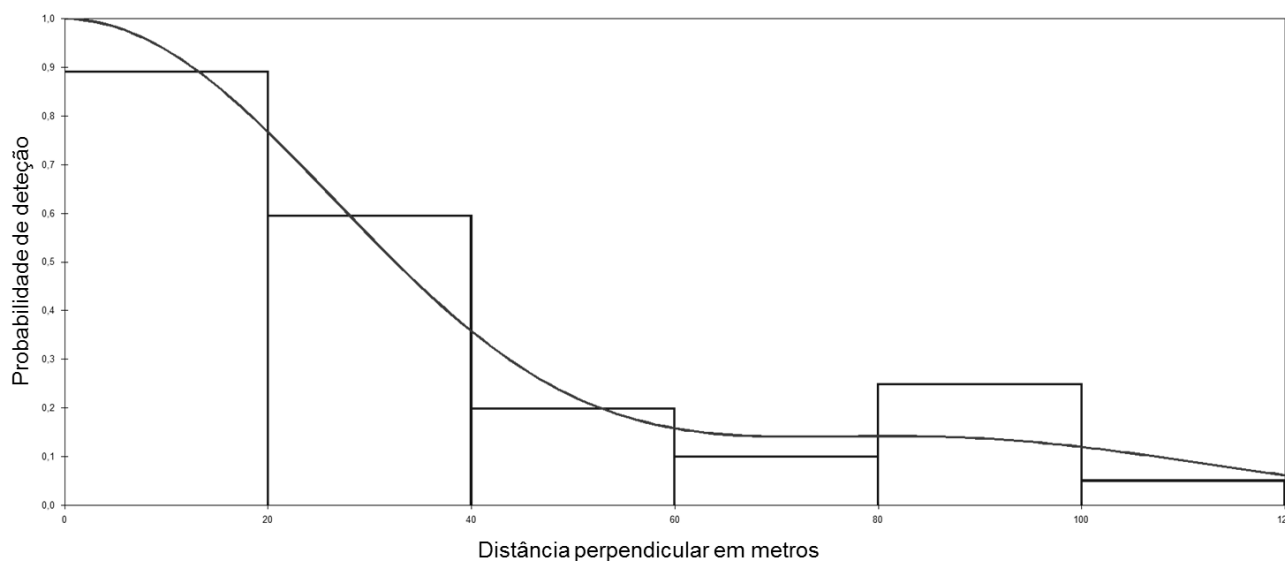


Figura 10 – Histograma que constitui a representação gráfica da função de detecção para a ZCT da Cubeira ilustrando a probabilidade de detectar veados em função da distância perpendicular.

A densidade populacional para a ZCT da Cubeira no período de transição Inverno-Primavera foi estimada em 23,49 veados/100ha, sendo que a variação, para um intervalo de confiança de 95%, situou-se entre os 15,83 e os 34,85 veados/100 hectares. Tendo em consideração estes valores, o número aproximado de animais na ZCT da Cubeira para este período do ano foi de 366 (247 – 544).

Os valores de densidade obtidos patenteiam uma discrepância acentuada entre as populações estudadas e poderá constituir parte da justificação para o enviesamento do rácio sexual favorável às fêmeas que, como constatável, se acentua com o aumento da densidade (Kruuk *et al.* 1999). Relativamente a este indicador, a realização de um teste de proporção contra valores fixos demonstrou existem diferenças significativas nas proporções machos/fêmeas das três populações estudadas ($p < 0,05$). Algumas razões alternativas e/ou complementares e ressalvas no que concerne ao indicador relativo ao rácio sexual foram mencionadas anteriormente. Ao nível da produtividade não se observam diferenças significativas entre as populações podendo ser considerados valores aceitáveis tendo em consideração o panorama europeu.

Na Figura 11 é apresentada a distribuição espacial da espécie pelas áreas de estudo e respectivas variações de densidade.

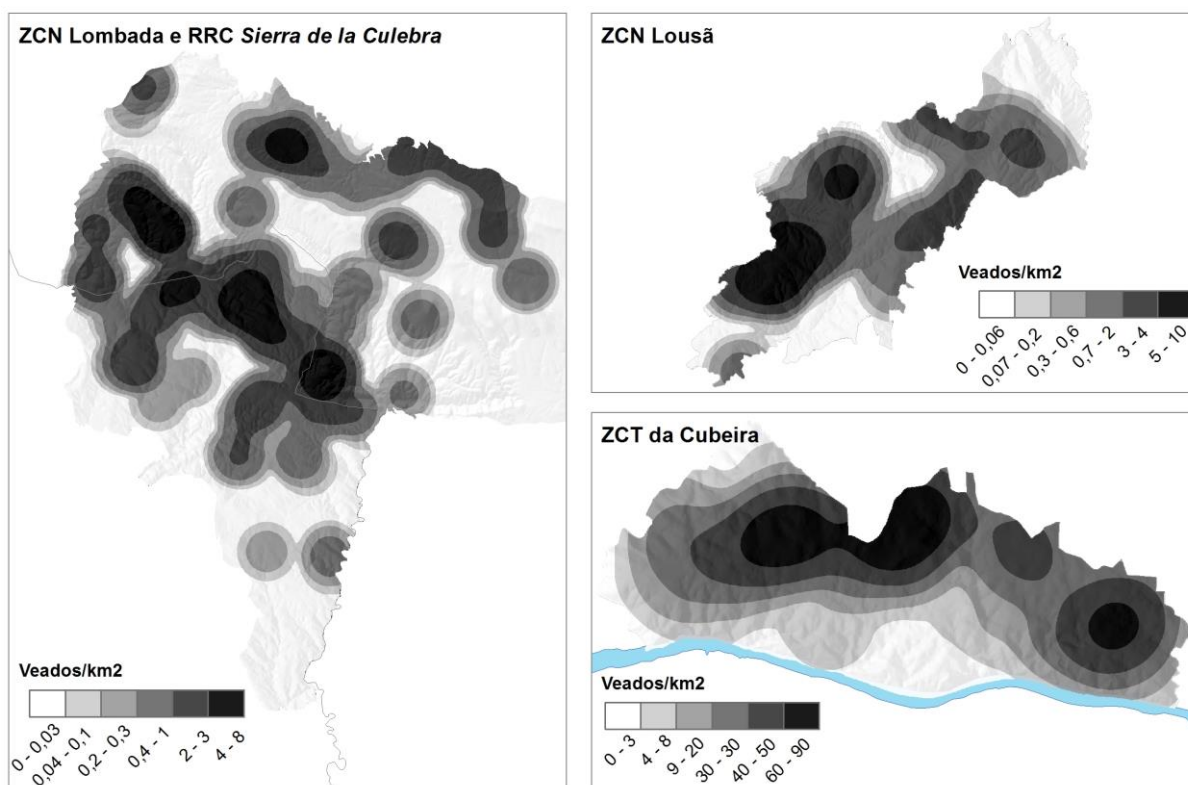


Figura 11 – Variação da densidade de veados nas áreas de estudo obtidas com recurso ao *Kernel Density Estimator* do software *ArcGIS 10.1*.

Organização social

A composição e tamanho dos grupos são duas características inerentes às unidades sociais que variam ao longo de um ciclo anual por razões distintas. A composição relaciona-se com a fenologia e interações sociais da espécie, enquanto o tamanho reflete as pressões ambientais às quais a mesma está sujeita.

Na Figura 12 é apresentada a composição dos grupos contemplando os dados compilados de todas as épocas de amostragem no período de acasalamento para as três populações.

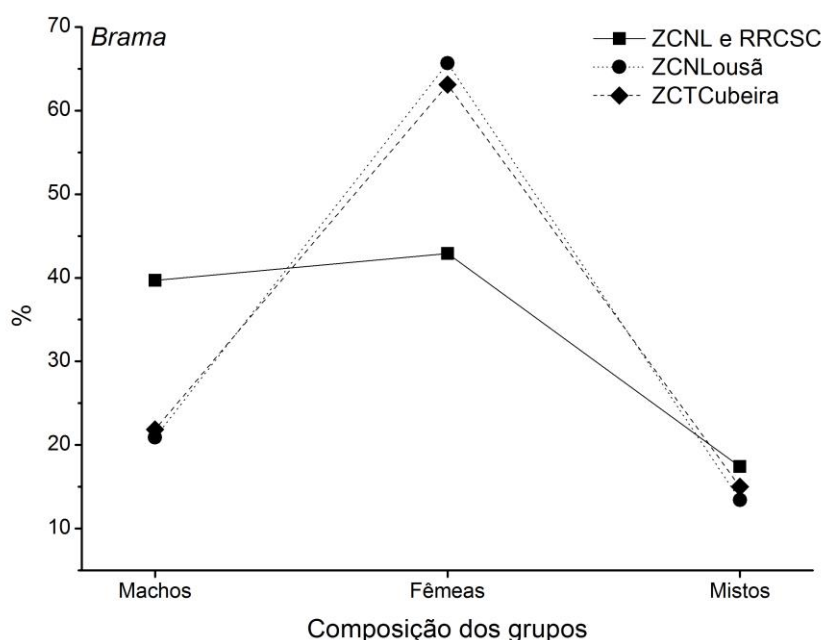


Figura 12 – Composição dos grupos na época da brama para as três populações estudadas.

À semelhança do observado para a estrutura populacional, a composição dos grupos é elucidativa do seu equilíbrio. A população do Noroeste peninsular apresenta um rácio no que concerne aos grupos de machos e fêmeas mais próximo da unidade do que as restantes populações. Relativamente aos machos, para este período, nesta população, os solitários representaram 78% deste grupo o que é compreensível face ao comportamento territorial dos machos defensores de um harém. Atendendo ao total populacional, os machos solitários representaram 36% deste. Apesar das discrepâncias

relativas à composição dos grupos, a percentagem de machos solitários na ZCN da Lousã foi igualmente considerável (79% considerando o grupo e 23% considerando o total populacional). Para a ZCT da Cubeira, os machos solitários representam 94% do grupo e 22% do total da população.

O efeito da densidade populacional no rácio sexual tem sérias implicações na gestão de uma população. O controlo da densidade tem que ser realizado dependendo dos objetivos que se pretendam alcançar, ganhando relevância quando o pretendido é maximizar a produção de machos. Uma densidade populacional adequada à capacidade de carga do meio contribui para uma estrutura populacional equilibrada o que possibilita uma exploração sustentável a longo prazo.

O conhecimento do tamanho e organização dos grupos (organização social), assim como das interações que ocorrem no seio dos mesmos (estrutura social) é essencial no sentido de descrever a dinâmica populacional de uma espécie. A organização e estrutura social relacionam-se com as táticas antipredatórias, comportamento alimentar e reprodutivo, uso do *habitat* (Jarman 1974), além de que, as relações de dominância afetam a resposta reprodutiva do veado (ver Goodwin *et al.* 1999).

A Figura 13 representa a distribuição dos tamanhos das unidades sociais para a mesma época e para as mesmas três populações. Os grupos apresentam tamanhos enviesados reflectindo a tendência descrita por Krause & Ruxton (2002), onde se afirma que a frequência dos grupos pequenos é mais elevada.

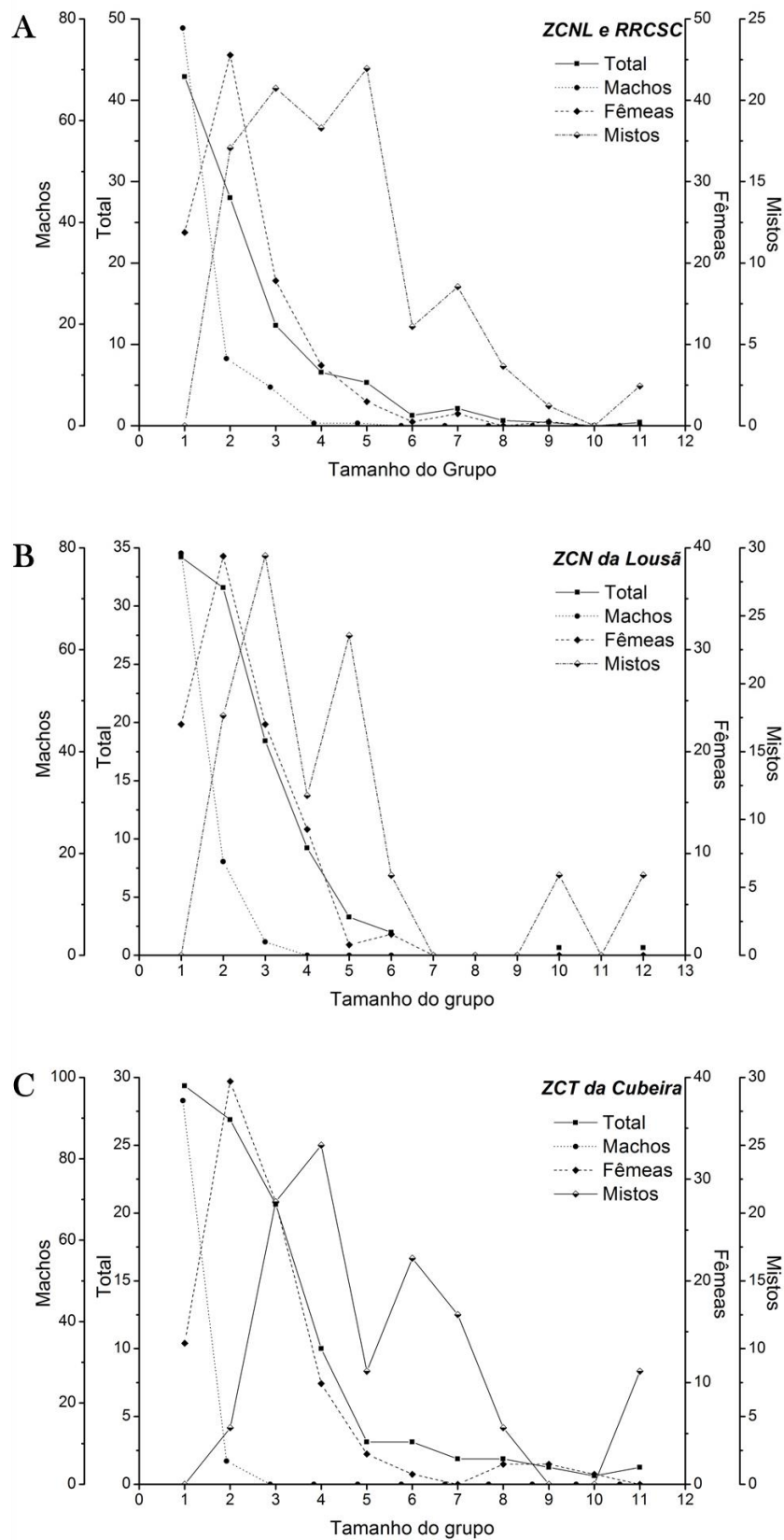


Figura 13 – Distribuição do tamanho das unidades sociais para a época da brama para as três populações estudadas. A) ZCN da Lombada e RRC *Sierra de la Culebra*; B) ZCN da Lousã e C) ZCT da Cubeira.

Na Tabela 5 são apresentados os valores gerais dos tamanhos dos grupos no período de acasalamento para as três populações. Além do tamanho médio do grupo, é apresentada informação relativa ao tamanho médio do grupo experienciado individualmente.

Tabela 5 – Tamanho médio dos grupos e tamanho médio do grupo experienciado individualmente, com os respetivos intervalos de confiança a 95%, para as três populações na época da brama.

População	Método	Total	Machos	Fêmeas	Mistos
ZCN Lombada e RRC Sierra de la Culebra	Grupo	2,25	1,32	2,31	4,35
		(2,11 - 2,40)	(1,23 - 1,41)	(2,14 - 2,49)	(3,96 - 4,77)
	Crowding	3,47	1,67	2,99	5,16
		(3,16 - 3,82)	(1,50 - 1,90)	(2,70 - 3,40)	(4,70 - 5,83)
ZCN Lousã	Grupo	2,32	1,24	2,36	4,53
		(2,09 – 2,59)	(1,11 – 1,39)	(2,15 – 2,58)	(3,53 – 6,29)
	Crowding	3,39	1,43	2,89	6,09
		(2,88 – 4,42)	(1,19 -1,73)	(2,63 – 3,23)	(4,17 – 8,68)
ZCTC	Grupo	2,79	1,06	2,80	5,25
		(2,49 - 3,13)	(1,00 – 1,14)	(2,50 – 3,16)	(4,46- 6,29)
	Crowding	4,30	1,11	3,84	6,29
		(3,72 – 5,06)	(1,00 – 1,25)	(3,24 – 4,73)	(5,23 – 7,82)

O tamanho médio dos grupos no período de acasalamento é idêntico nas três populações estudadas, apesar dos valores de densidade destas serem consideravelmente diferentes. De facto, o tamanho dos grupos resulta de uma série de condicionantes ambientais distintas constituindo um compromisso entre custos e benefícios.

Os dados apresentados dizem respeito ao período da brama contudo, os vários tipos de grupos registam um aumento progressivo do tamanho desde o fim do verão até à primavera do ano seguinte. Os grupos mistos diluem-se, formando-se unidades matriarcais e grupos de machos (Figura 14). Esta alteração em número de indivíduos por grupo é uma resposta às variações dos efeitos da predação, quando presente, rendimento/disponibilidade alimentar e uso do *habitat* (Pulliam & Caraco 1984).

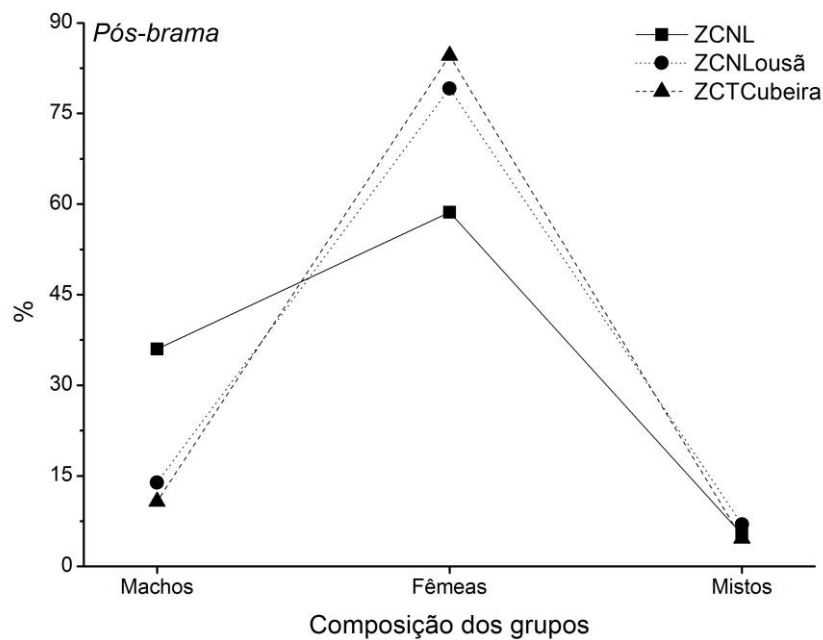


Figura 14 – Composição dos grupos na época posterior à brama para as três populações estudadas.

Regista-se um decréscimo expectável do número de grupos mistos observados. O decréscimo da frequência relativa dos grupos de machos é mais evidente na população do Noroeste Peninsular. Este facto resulta de que, na época de reprodução, se terem observado muitos machos jovens solitários que foram afastados dos haréns, o que aumentou a frequência. Terminada esta época, os machos formam grupos unissexuais mais numerosos (ver Figura 15 e Tabela 6). Sustentando esta informação importa salientar que para a população do Noroeste da Península Ibérica, a frequência relativa de machos solitários desceu para 38% e 14% considerando o grupo e o total populacional, respetivamente. A percentagem de grupos matriarcais permanece elevada. Para a ZCN da Lousã este valor cifrou-se, para o grupo, nos 50% e para o total, nos 12%. Por fim, na ZCT da Cubeira os valores percentuais registaram 43%, para o grupo, e 6,5%, para o total populacional.

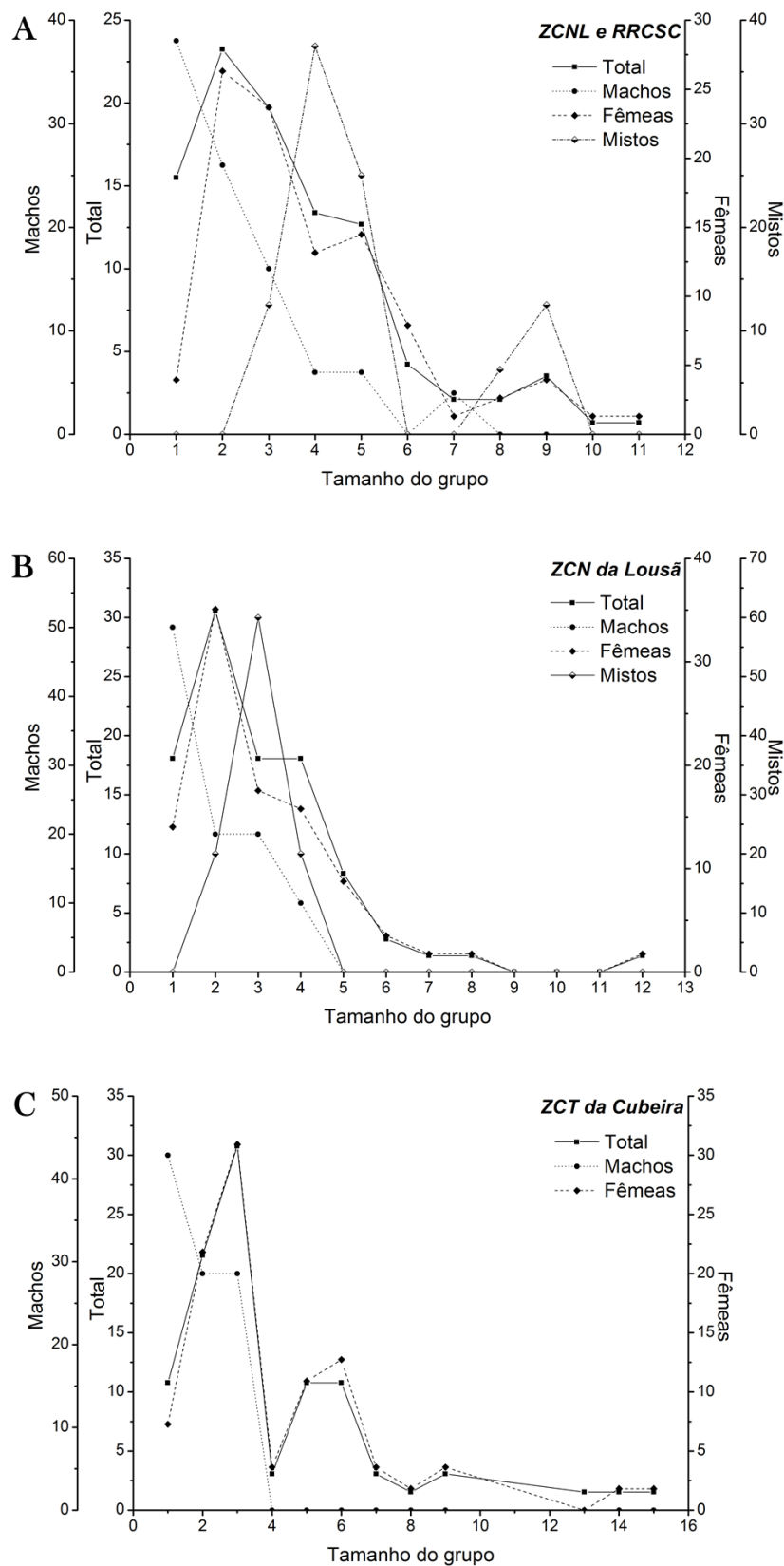


Figura 15 – Distribuição do tamanho das unidades sociais na época posterior à brama para as três populações estudadas. A) ZCN da Lombada e RRC *Sierra de la Culebra*; B) ZCN da Lousã e C) ZCT da Cubeira.

Na Tabela 6 são apresentados os valores gerais dos tamanhos dos grupos no período de pós-acasalamento para as três populações.

Tabela 6 – Tamanho médio dos grupos e tamanho médio do grupo experienciado individualmente, com os respetivos intervalos a 95%, para as três populações na época posterior à brama.

População	Método	Total	Machos	Fêmeas	Mistos
ZCN Lombada e RRC Sierra de la Culebra	Grupo	3,72	2,29	3,93	5,56
		(3,33 – 4,18)	(1,90 – 2,75)	(3,58 – 4,42)	(4,50 – 7,06)
	Crowding	5,62	3,31	5,14	6,84
		(4,83 – 6,88)	(2,69 – 4,19)	(4,49 – 6,03)	(5,24 – 9,37)
ZCN Lousã	Grupo	3,01	1,90	3,12	4,00
		(2,64 – 3,50)	(1,30 – 2,50)	(2,68 – 3,72)	(3,20 – 4,40)
	Crowding	4,17	2,47	4,36	4,10
		(3,51 – 5,71)	(1,71 – 3,28)	(3,57 – 6,08)	(3,47 – 4,65)
ZCTC	Grupo	4,03	1,86	4,15	7,00
		(3,40 - 4,85)	(1,14 - 2,43)	(3,47 - 5,02)	(3,00 - 10,3)
	Crowding	6,17	2,23	6,07	9,67
		(4,91 - 8,12)	(1,44 - 2,76)	(4,85 - 8,13)	(3 - 11,97)

É possível observar um aumento generalizado do tamanho dos grupos do período de acasalamento para o período fenológico seguinte, refletindo o já descrito para outras populações em contexto mediterrânico.

O tamanho das unidades sociais é um reflexo das pressões ambientais às quais a espécie está sujeita, sendo que, a plasticidade do veado permite-lhe selecionar sazonalmente o *habitat* de acordo com a disponibilidade alimentar (Carranza & Valencia 1992).

Em ambientes mediterrânicos, os verões secos e quentes contrastam com os invernos suaves. Este aspeto reflete-se na produção de herbáceas que, nos ecossistemas mediterrânicos, atinge o seu máximo na primavera e fins de outono, prolongando-se por todo o período invernal. O verão constitui um período de escassez. Deste modo, os indivíduos da espécie selecionam os espaços abertos preferencialmente durante o inverno (ver Carranza *et al.* 1991). Em ecossistemas mediterrânicos, o maior

tamanho dos grupos resulta da exploração comum de pradarias ricas em plantas herbáceas, designadamente gramíneas e leguminosas. Neste período, as herbáceas são o constituinte maioritário da dieta do veado (Rodriguez-Berrocal 1978).

De acordo com Carranza & Valencia (1992), o maior tamanho dos grupos no inverno-primavera deve ser explicado com base no uso frequente das áreas abertas nessas épocas, contrariamente, ao uso de *habitats* fechados como bosques e matos bem desenvolvidos durante o verão. Nas populações mais setentrionais, a sazonalidade no uso do *habitat* e, conseqüentemente, tamanho dos grupos, registam uma tendência inversa (ver Clutton-Brock *et al.* 1982; Loe *et al.* 2005). Estes factores justificam, em parte, o aumento dos tamanhos dos grupos de um período para o outro.

A população da ZCN da Lombada e *RRC Sierra de la Culebra* está sujeita à predação por parte do lobo-ibérico. Potencialmente, este aspeto assume extrema relevância no tamanho médio e organização dos grupos. De facto, as táticas antipredatórias dos ungulados baseiam-se na ocultação quando se alimentam em *habitats* fechados e na vigilância coletiva quando as áreas abertas são as mais exploradas, o que se repercute no aumento do tamanho dos grupos nestas últimas (Jarman 1974). Além da vigilância, o efeito de dissolução e “confusão”, são duas das vantagens da formação de grupos maiores (ver Senar 1994). Em áreas abertas, quanto maior o grupo, maior a efetividade da proteção contra a predação. A proteção é maior nas posições centrais do grupo (Treisman 1975), sendo esses, os locais ocupados pelas crias e animais mais jovens num grupo matriarcal. No centro, a eficiência da alimentação é mais elevada, uma vez que na periferia a vigilância tem que ser efectuada continuamente (Robinson 1981 *in* Carranza & Reyna 1987). A maior eficácia da vigilância de um grupo relativamente a um único indivíduo aumenta a percentagem de tempo dedicado exclusivamente à alimentação (Lima & Dill 1990). No inverno, em determinadas áreas, é possível observar grupos com algumas dezenas de animais. Estes grupos, na realidade, não representam uma unidade social, mas sim um conjunto de indivíduos formado por vários grupos familiares que se associam nas melhores áreas de alimentação.

Referências bibliográficas

- Akaike H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 19: 716-723.
- Barroso I. & Rosa J.L. (1999). O Veado. Coleção Património Natural Transmontano. João Azevedo Editor. Mirandela. 74pp.
- Carranza J. & Reyna L.A. (1987). Spatial Organization of female groups in red deer (*Cervus elaphus* L.). *Behavioural Processes*. 14: 125–135.
- Carranza J., Hidalgo de Trucios S.J., Medina R., Valencia J. & Delgado J. (1991). Space use by red deer in a Mediterranean ecosystem as determined by radio-tracking. *Applied Animal Behaviour Science*. 30: 363–371.
- Carranza J. & Valencia J. (1992). Organización social del ciervo en hábitat mediterráneo. *Miscellanea Zoologica*. 16: 223–232.
- Carvalho J. (2011). Monitorização da população de veado (*Cervus elaphus* L.) na Zona de Caça Nacional da Lombada, Nordeste Transmontano: Época 2010/2011. Relatório de Estágio para obtenção da licenciatura em Biologia. Universidade de Aveiro.
- Clutton-Brock T.H., Guinness F.E. & Albon S.D. (1982). Red deer: Behavior and Ecology of two sexes. Edinburgh University Press. Edinburgh. 400pp.
- Clutton-Brock T.H., Albon S.D. & Guinness F.E. (1984). Maternal dominance, breeding success and birth sex ratios in red deer. *Nature*. 308: 358–360.
- Goodwin N., Hayssen V., Deakin D.W. & Flint A.P.F (1999). Influence of social status on ovarian function in farmed red deer (*Cervus elaphus*). *Physiology & Behavior*. 65: 691–696.
- Jarman P.J. (1974). The social organisation of antelope in relation to their ecology. *Behaviour*. 48: 215–267.
- Krause J. & Ruxton G.D. (2002). Living in groups. Oxford University Press. 228pp.

Kruuk L.E.B, Clutton-Brock T.H., Albon S.D., Pemberton J.M & Guinness F.E. (1999). Population density affects sex ratio variation in red deer. *Nature*. 399: 459–461.

Lima S.L & Dill L M. (1990). Behavioral decisions made under the risk of predation: a review and prospectus. *Canadian Journal of Zoology*. 68: 619–640.

Loe L.E., Bonenfant C., Mysterud A., Gaillard J.M., Langvatn R., Klein F., & Stenseth N.C. (2005). Climate predictability and breeding phenology in red deer: timing and synchrony of rutting and calving in Norway and France. *Journal of Animal Ecology*. 74(4): 579–588.

Pulliam H.R. & Caraco T. (1984). Living in groups: is there an optimal group size? Pp.122-147. Em: Krebs J.R. & Davies N.B. (eds). *Behavioural Ecology: An Evolutionary Approach*. Blackwell Scientific Publications. Oxford. 464pp.

Rodríguez-Berrocal J. (1978). Introducción al estudio y valoración de los recursos forestales y arbustivos para el ciervo en el área ecológica de Sierra Morena. I Estudio de la dieta del ciervo. *Archivos de Zootecnia*. 27: 73–82.

Santos J. (2009). Estudo populacional do veado (*Cervus elaphus* L.) no Nordeste Transmontano. Tese de Mestrado em Ecologia, Biodiversidade e Gestão de Ecossistemas. Universidade de Aveiro.

Senar J.C. (1994). Vivir y convivir: La vida en grupos sociales. Pp.205-233. Em: Carranza J. (eds.). *Etología: Introducción a la Ciencia del Comportamiento*. Publicaciones de la Universidad de Extremadura. Cáceres. 591pp.

Thomas L., Laake J.L., Rexstad E., Strindberg, S., Marques, F.F.C., Buckland S.T., Borchers D.L., Anderson D.R., Burnham K.P., Burt M.L., Hedley S.L., Pollard J.H., Bishop J.R.B. and Marques T.A. (2009). Distance 6.0. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. <http://www.ruwpa.st-and.ac.uk/distance/>

Treisman M. (1975). Predation and the evolution of gregariousness. I. Models for concealment and for evasion. *Animal Behaviour*. 23: 779–800.

Trivers R.L. & Willard D.E. (1973). Natural selection of parental ability to vary the sex ratio of offspring. *Science*. 179: 90–92.

Capítulo 4

Habitat suitability model for red deer (*Cervus elaphus* Linnaeus, 1758):
spatial multi-criteria analysis with GIS application

Carvalho J., Martins L., Silva J.P., Santos J., Torres R.T. & Fonseca C. (2012). Habitat suitability model for red deer (*Cervus elaphus* Linnaeus, 1758): spatial multi-criteria analysis with GIS application. *Galemys*, 24: 47 – 56.

O artigo é transcrito na íntegra correspondendo à versão final publicada.

Habitat suitability model for red deer (*Cervus elaphus* Linnaeus, 1758): spatial multi-criteria analysis with GIS application

Monitoring population trends is essential in wildlife management. In fact, to identify those environmental conditions linked to habitat use and selection is a key task in any conservation plan. Our aim in this work was predict red deer (*Cervus elaphus* Linnaeus, 1758) colonization in the Lombada National Hunting Area (LNHA), one of the largest wild populations in Portugal. The sharply increase in density and range of red deer over the last decades justify the careful monitoring of this population. A multi-criteria model using the Analytical Hierarchic Process (AHP) and GIS weighted spatial analysis with matrix algebra techniques were applied in this approach. This method was developed in three distinct phases: i) setting the objectives, ii) analysing pairs of factors through the application of a comparison matrix and iii) its synthesis with the weight assignment that followed a predetermined numeric range according to the ecological requirements of the study species. The variables used in the red deer habitat suitability model included the hydrographic network, asphalted roads, relief aspects and land use. From October 2010 to March 2011, a total of 129 observations along line transects were recorded and used to validate the final model. By using the AUC method it was obtained a success rate of 72%. Our results suggest that this method would be applied on a larger scale being suitable to predict red deer expansion. To our knowledge, this is the first study performed in Portugal using this methodological approach to assess the red deer-habitat relationships.

Keywords: *Cervus elaphus*, GIS, habitat suitability, multi-criteria models.

Capítulo 4

Introduction

Understanding how species are spatial and temporally distributed is a central topic in ecology. In fact, to identify those environmental conditions linked to habitat use and selection is a key task in any conservation plan (Guisan & Zimmermann 2000). Species distribution models (SDM) are currently recognised as extremely effective tools providing valuable and quantitative information by displaying the most important conditions and resources required by the individuals in a given spatial context (Guisan & Thuiller 2005). Therefore, habitat suitability models are important tools in conservation planning and game management (Pearce & Ferrier 2000). Even though predicting natural processes is difficult (e.g. Radeloff *et al.* 1999, Shi *et al.* 2006), many methodologies and mathematical approaches have been developed for such purposes. However, variability registered from different modelling techniques is one of the main issues for species distribution assessment. In this line, Araújo & New (2007) suggests an ensemble approach, where several models were tested in order to improve the results reliability. Nevertheless, the strategy to follow should reflect a trade-off between the specific objectives and the available data. There is a broad range of modelling techniques available to explore the correlation between species occurrence and predictor variables (e.g. food distribution, cover, human disturbances) (Guisan & Zimmermann 2000). These techniques include logistic multipleregression (Pearce & Ferrier 2000), regression trees (Stankovski *et al.* 1998), geographically weighted regression (Shi *et al.* 2006), spatial and space-time correlations (Epperson 2000) and some of the most widespread techniques as Maxent (Phillips *et al.* 2006) and BIOCLIM (Beaumont *et al.* 2005).

Over the last three decades, red deer (*Cervus elaphus*) populations have increased in numbers and distribution in Portugal (Vingada *et al.* 2010). A wide range of factors are related with the widespread distribution of this species, namely natural dispersion (Vingada *et al.* 2010), decrease of human pressure and rural exodus with consequent abandonment of agricultural fields. Since 1970's, red deer populations have recovered from a situation of virtual extinction being now common in many regions of Portugal, but especially close to the Spanish border (Vingada *et al.* 2010). Red deer play an important ecological role in maintaining floristic diversity and seed dispersal (Malo & Suárez 1995, Schütz *et al.* 2003). Moreover, it is also a wild prey of the endangered

Iberian wolf (*Canis lupus signatus* Cabrera, 1907). In Portugal, one of the main red deer populations is located in the northeast of the country (Trás-os-Montes region). Santos (2009) showed that density varied between 2.27 and 4.70 red deer/km², however Carvalho (2011) reported a density decrease to 1.07 - 2.87 red deer/km² in the same population. By applying indirect methods (pellet group counts), Torres *et al.* (2012) identified the environmental variables that affect the distribution and occurrence of this population, showing that red deer was found in areas further away from roads, urban settlements, agricultural fields selecting patches with medium ground cover of shrubs and sparse tree cover.

In this study, it was applied an Analytical Hierarchy Process (AHP) (Saaty & Vargas 1991) coupled with Geographic Information Systems (GIS). This is the first study performed in Portugal using this methodological approach to analyse red deer – habitat relationships.

Due to current increase of red deer in Portugal, the identification of the environmental factors potentially linked to further expansions is needed. Hence, this study have two main mains: i) to construct a habitat suitability map for red deer in *Lombada* National Hunting Area, incorporating important ecological variables for the species and ii) highlight the use of a multi-criteria decision making based on the Analytical Hierarchy Process as an alternative way to define suitability classes.

Material and Methods

Study area

The study was carried out in the Lombada National Hunting Area (20.830ha), Bragança (Figure 16). Most of the study area is located inside the Montesinho Natural Park that is part of the European Union's Natura 2000 Network. The altitude ranges from 530 to 1073 m.a.s.l. The climate, mainly Mediterranean, is influenced by three chains of mountains and molded by Castellano-Leones plateau. Mean annual temperature varies between 8 and 12.5°C, registering a mean temperature of 3°C and 21°C in coldest and warmest month, respectively (Aguar 2000). The vegetation, although dominated by scrubland, is varied and characterized by pyrenean oak (*Quercus pyrenaica*), sweet chestnut (*Castanea sativa*), scots pine (*Pinus sylvestris*), european black pine (*Pinus nigra*), maritime pine (*Pinus pinaster*), and holm oak (*Quercus rotundifolia*). Main understory species are *Erica australis*, *Cystisus* sp., *Pterospartum tridentatum*, *Halimium lasianthum*

subsp. *alyssoides* and *Cistus ladanifer*. Scattered pastures and agricultural crops can still be found along the study area. Three main rivers and some streams cross the study area. The riparian vegetation associated is mostly composed by common alder (*Alnus glutinosa*), ash (*Fraxinus angustifolia*), black poplar (*Populus nigra*) and white willow (*Salix salviifolia*). Two others ungulates are present: roe deer *Capreolus capreolus* (Linnaeus, 1758) and wild boar *Sus scrofa* Linnaeus, 1758. Iberian wolf is one the red deer main predators and densities have been calculated in 1.6 - 3.1 wolves/100 km² (Moreira *et al.* 1997). Taking into account the European context, where wolf densities are generally <1 wolf/100 km² (Boitani 2000), these values can be considered high. Human population density is low (9.5 people/km²). Road network can also be considered low since there are a few national roads connecting the small villages. Red deer hunting is allowed in the study area and hunting bags vary between 3-5 per year.

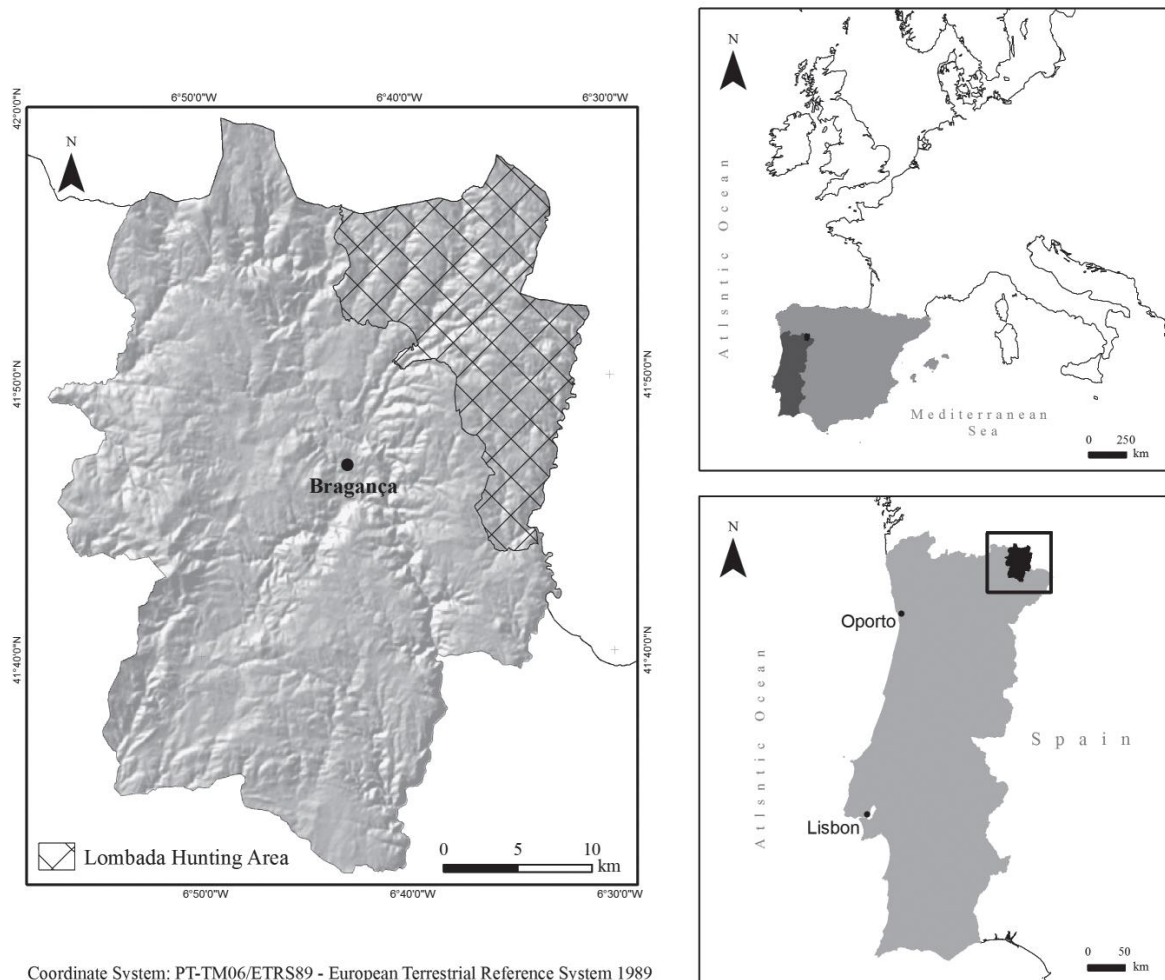


Figure 16 – Location of study area.

Survey design

Field work was conducted between October 2010 and March 2011 using line transects placed systematically to provide equal coverage of habitats that occur in the area (Figure 17). With this, the bias associated with the systematic prospection of areas with high or low deer densities was reduced (Marques *et al.* 2001). The sampling comprised two distinct periods: mating season (September - October) and late winter. Twenty one lines transects were established, with an average length of 4.3 (1.78 - 6.82) km and totalling 90.3 km. Short transects gave us the possibility to achieve a better representation of whole study area. When a red deer or a group was detected, the distance from the observation point to the animal group was recorded. Through GPS location and trigonometric operations we could estimate the exact position of the animals observed that was then used to validate the final model.

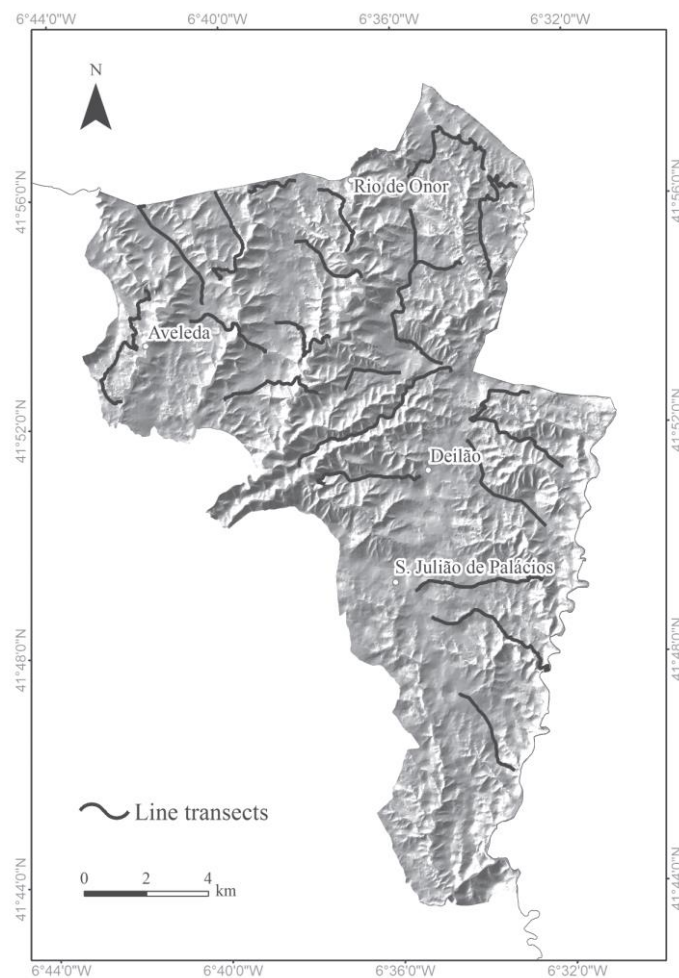


Figure 17 – Line transects in the study area.

Environmental variables

The variables were selected considering their applicability to the scale of our study area, relevant predictive power (Post & Forchhammer 2002) and their relationship with the ecological requirements of the species (Guisan & Zimmermann 2000). Based on previous studies (San-José *et al.* 1997, Torres *et al.* 2012) and in our own field experience, we selected the variables that could potentially influence the species occurrence. These variables were classified into four classes: habitat composition (Thomas *et al.* 1979, San-José *et al.* 1997, Debeljak *et al.* 2001), human disturbance (Hewinson *et al.* 2001, Dziba *et al.* 2003), topographic factors (Acevedo *et al.* 2011) and distance to water courses (Thapaliya 2008). For the habitat composition class we have used three general land use predictor variables divided in forest, agricultural and urban areas. This information was obtained from CORINE Land Use/Land Cover database (CLC06) with spatial resolution (pixel width) of 250 meters. Forest areas include the broadleaved and coniferous woodland, the wetlands and the scrubland (dense and sparse). Agricultural land included herbaceous and woody crops, the arable horticulture and the heterogeneous crops and urban areas included houses, buildings and industrial areas. Since red deer hunting is allowed in the study area, it is expected that human disturbance factors will influence red deer distribution as they can be considered as analogues to predation risk (Lima & Dill 1990). Therefore, we consider the distance to the closest urban area and to the closest asphalted road (Acevedo & Cassinello 2009). Additionally, we selected a topographic factor that can potentially affect the red deer occurrence, *i.e.* the relief aspect (Acevedo *et al.* 2011). This variable was classified in five discrete classes (North, West, South, East and Flat).

Multi-criteria analysis: an overview of the AHP method

The criteria weights were obtained using the Analytical Hierarchy Process (AHP) (Saaty & Vargas 1991) and employed in the GIS based MCDM. A MCDM combines the environmental factors under analysis in a single parameter of evaluation (Chen *et al.* 2010). This method allows integrating heterogeneous data derived from different sources and correlating their respective weights. The main goal of this approach is to achieve the best solution that takes into account all the conflicting input criteria. The AHP process is based in three fundamental steps: i) develop the AHP Hierarchy, ii) pairwise comparisons between all factors using a given scale, and iii) final weights definition. The final solution

will be a choice between alternatives or criteria, which should be measured and evaluated.

A weighted linear combination was used since it is the most prevalent procedure for multi-criteria evaluation. Factors (e.g. variables) are combined together by applying a weight to each one followed by the sum of the weights applied to each factor. The output is a habitat suitability map.

$$S = \sum (w_i x_i)$$

(where S is suitability, w_i is the weight of factor i , x_i is the criterion score of factor i).

In this particular case the final result is a Habitat Suitability Index (HSI) that represents quantitatively the capacity of an area to fulfill the requirements of a study species.

Pairwise comparisons

Criteria weights were assigned according to Saaty's pairwise comparisons (Saaty & Vargas 1991). By using this comparison matrix subjectivity is minimized. In the first step, a qualitative numerical scale was used to score each pairwise comparison between the chosen criteria. Applying a scale using 9-point rating scale ranging from 1 (equal importance) to 9 (strongly more important) and reciprocal values, the relative preference between two variables under analysis was achieved (Table 7). Before any map algebra operations, all continuous variables were standardized (linear scaling and scale inversions) to avoid the effect of different measurements scales and to facilitate direct comparison.

$$x_i = \frac{(R_i - R_{min})}{(R_{max} - R_{min})}$$

(where R_i is the raw score of factor i).

The comparison matrix was then filled in both directions (Table 8). By using the given values, the specific weights for each criterion were calculated in order to be used in the weighted linear combination (Table 9).

Table 7 – Saaty’s pairwise comparisons. The values varies between 1 and 9 (factor on vertical axes is more important than the factor on horizontal axes) or 1/3 and 1/9 (factor on vertical axes is less important than the factor on horizontal axes).

Degree of importance	Definition
1	Equal importance
3	Weak importance
5	Strong importance
7	Very strong importance
9	Extreme importance
2, 4, 6, 8	Intermediate values
1/2, 1/3, 1/4, 1/5, 1/6, 1/7, 1/8, 1/9	Reciprocal values

Table 8 – Comparison matrix with the relative weight assigned to the factors under analysis.

Variables	Forest areas	Agricultural areas	Distance to asphalted road network	Distance to hydrographic network	Distance to urban areas	Relief aspects
Forest areas	1	-	-	-	-	
Agricultural areas	1/7	1	-	-	-	-
Distance to asphalted road network	1/5	1/3	1	-	-	-
Distance to hydrographic network	1/3	3	5	1	-	-
Distance to urban areas	1/9	1/5	5	1/7	1	-
Relief aspects	1/5	1/3	3	1/5	1/5	1

Table 9 – Final factor weights to be applied.

Forest areas	Agricultural areas	Distance to asphalted road network	Distance to hydrographic network	Distance to urban areas	Relief aspects
0.38	0.12	0.04	0.22	0.14	0.10

Data analysis

Based on the Area Under the Curve (AUC) method, the spatial model obtained was validated using the presence points of the red deer observed groups and pseudo-absences obtained through random points dispersed across all the study area. In the first time, a Receiver Operating Characteristics (ROC) plot was obtained by plotting true positive fraction on the y axis against false positive fraction on the x axis (Fielding & Bell 1997). According to the same authors, this method provides a way to compare classifiers testing the model accuracy and allowing their validation independent of distortions and potential bias. The output values vary between 0.5 and 1. The first indicates that the scores of the two groups do not differ, while the second expresses a non-overlapping of the group scores. Furthermore, by using a logistic regression we compared the predictive performance of the heuristic and statistical procedures. We evaluated the effectiveness of a multicriteria decision in order to test it as an alternative way to obtain habitat suitability maps in the lack of presence/absence data. The AUC value was used to measure the predictive performance of the models. The final model representing the habitat suitability for the study area was divided into three classes of suitability through the application of Natural Breaks method, which is one of the most common procedures to classify the quantitative data (Conolly & Lake 2006). As Hirzel *et al.* (2006) argued the low number of classes provides most honest and robust information. The method mentioned above is the most suitable for continuous data creating a set of logical and coherent classes through the definition of natural clusters (Jenks & Caspall 1971). The class ranges are defined comparing them with the distribution of the entire dataset allowing the identification of break points. Finally, the data available are divided in order to maximize the differences among the number of classes desired. Habitat evaluation and the selection of habitat patches with high suitability were performed through the comparison between high habitat suitability patches areas and the species annual home range. Since there was no information regarding red deer home ranges in the study area, we used home range values from other areas from the Iberian Peninsula that shared environmental characteristics with the study area.

Results

The habitat suitability model derived from weight assignment and matrix algebra operations corresponds with the observed densities of red deer in *Lombada* National Hunting Area (Figure 18). Our results show that the northern and central parts of study area are more suitable for red deer occurrence than the southern part. By dividing the suitability model into three quantitative parameters approximately 58% of the study area is classified as high suitable for red deer. The remaining 42% are divided into moderate (27%) and low suitability (15%). The comparisons of variables with the final model achieved makes possible to observe that red deer occurs in areas further away from roads, villages and agricultural fields. Moreover, they also show that red deer occurrence is mainly related with forested areas.

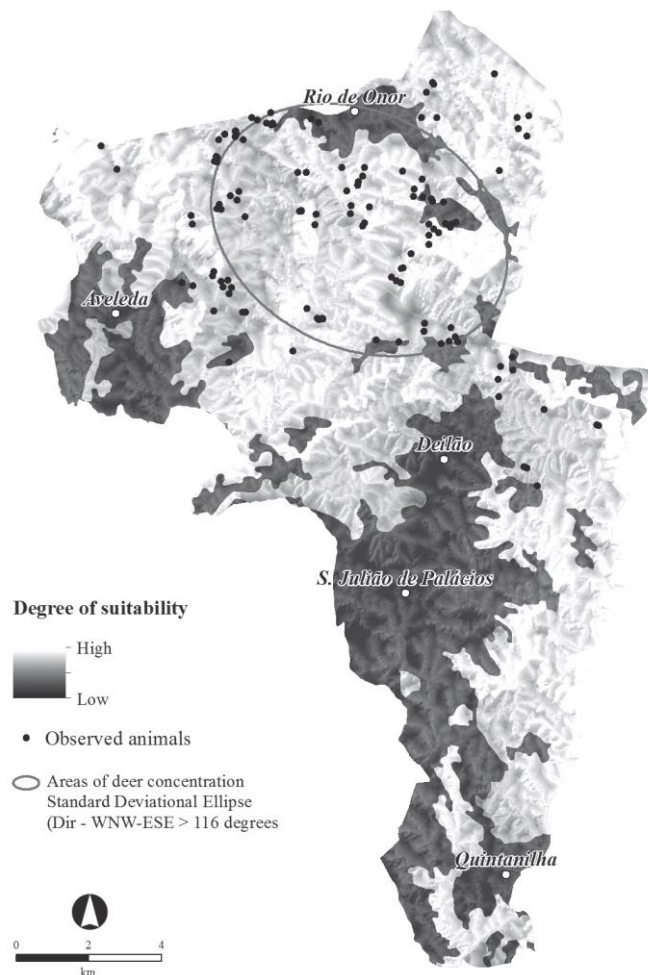


Figure 18 – Habitat suitability model output. Graduated colours from black (unsuitable areas) to white (suitable areas).

The presented model was validated with a total of 129 group observations. By applying the Area Under the Curve (AUC) method, it was obtained a success rate of 72% for the Analytical Hierarchic Process (Figure 19a). On the other hand, the value of the success rate obtained for the logistic regression model was 69% (Figure 19b). The values shown in Table 10 indicate the regression coefficients and the Wald statistic which tests the contribution of each predictor. By using the standard deviational ellipse in *ArcGIS* 9 3.1, it was determined an area of preferential red deer concentration (WNW - ESE) which corresponded to the northern part of the study area.

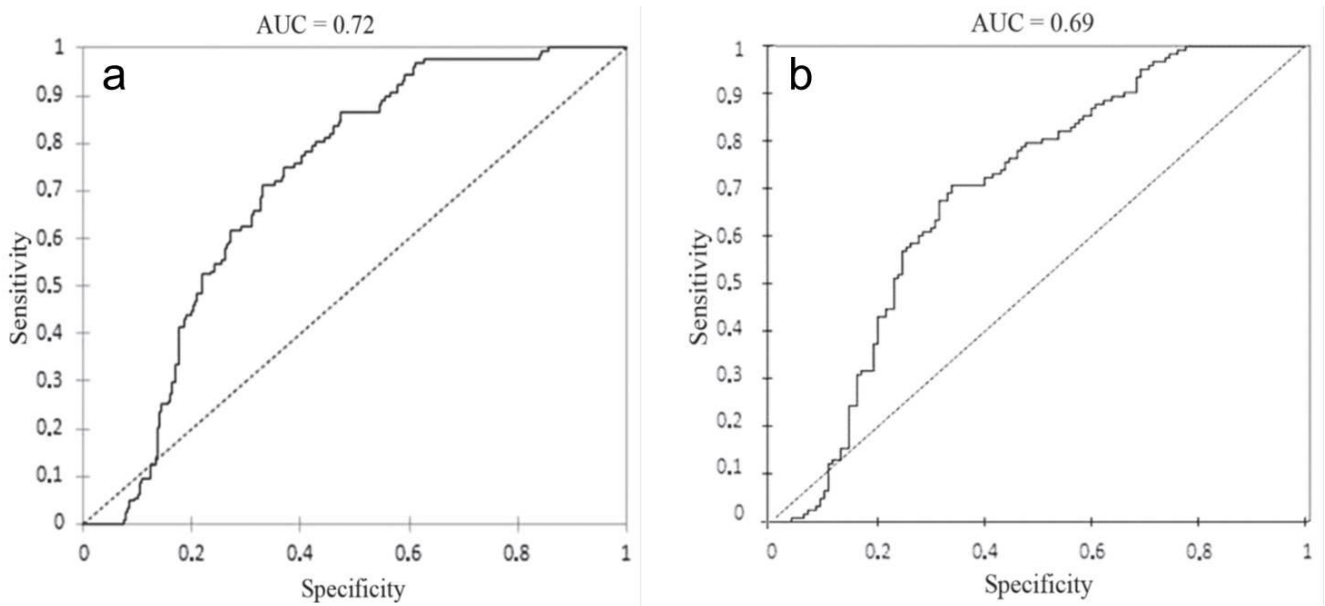


Figure 19 – Area under the ROC curve used to assess the accuracy of the models: **a)** Analytical Hierarchic Process, **b)** Logistic Regression.

Table 10 – Variables considered in the logistic regression model and their coefficients (β), standard deviation (SD), Wald test values and significance (p).

Variables	β	S.D.	Wald	p
Land use	0.003	0.004	0.754	0.385
Distance to asphalted road network	0.001	0.001	3.534	0.600
Distance to hydrographic network	0.001	0.001	4.310	0.042
Distance to urban areas	0.002	0.001	12.858	0.001
Relief aspects	-0.054	0.116	0.215	0.643
Constant	-2.244	1.241	3.269	0.071

Discussion

To our knowledge, this is the first study performed in Portugal using AHP and GIS weighted spatial analysis with matrix algebra techniques to assess the red deer - habitat relationships. Our model are in agreement and reflect in a spatially and qualitative way the results of previous studies conducted by Torres *et al.* (2012) for the same study area. Based on these authors' results, we have defined the weights to be applied and we hypothesized that the red deer occurrence might be mainly related with macro-habitat structure and human disturbance factors. Our results show that red deer used areas with coniferous and deciduous forests, where they can find refuge from adverse climatic conditions, human activities and wolf predation, interspersed by scrubland and meadows. These mixed areas with higher densities of heather and *Leguminosae* interspersed with ground cover and sparse tree cover provides food availability, which is essential to this species as it is considered a capital breeder (Clutton-Brock *et al.* 1983). Such ecotones provide a greater variety of plant species locally abundant. These transition habitats are important for red deer to cope with seasonal changes in resource availability by feeding grasses and/or shrubs (Putman & Flueck 2011) but they also provide hiding cover, important in our study area against wolf predation.

Our results showed that only 15% of the study area, the southwest part, comprises less suitable habitat for red deer. This might be due to the fact that this area is characterized by a fragmented landscape divided into small patches of agricultural crops with high pressure of human disturbance and livestock. Another important factor is the presence of feral dogs in the close vicinity of human settlements and the high pressure from illegal hunting, which is believed to be commonplace in the study area. According to Jiang *et al.* (2009), human activity is an important factor that influences red deer space use.

Mysterud *et al.* 2002 showed that agricultural crops provided an important food resource for red deer in winter. However, our results showed that red deer used areas further away from agricultural fields, which is in agreement with Torres *et al.* (2012). Interestingly, this fact deserves future investigations as this can be an adaptation to contrasting seasonality. In our study area, this fact can be related with the adoption of hiding behaviours to reduce the predation risk. Some studies demonstrated a preference for red deer to shrub species like *Pterospartum tridentatum*, *Cistus ladanifer*, *Halimium lasianthum* subsp. *alyssoides*, *Rubus ulmifolius* and *Erica* sp. (Alvarez & Ramos 1991, Ferreira 1998). In the northern part of our study area, these species predominate thereby suggesting that red deer fulfills its food requirements in the shrub layer rather than in the agricultural fields. Our results also show that some groups occurred in areas considered unsuitable scattered throughout the northern part of the study area. These areas correspond to small patches close to disturbance factors and close to agricultural crops and apparently do not have shelter habitats. The presence of contiguous suitable habitat in their surrounding might explain the occurrence of red deer groups in these unsuitable areas.

Allen *et al.* (1987) stated that a habitat suitability model must consider the concept of "evaluation unit". According to the same authors, this is defined as the minimum area of contiguous habitat that a population requires to persist in each stage of its life cycle. Therefore, the habitat evaluation was made on the basis of animal annual home range. Since in our study area we do not have red deer home range information, we used values recorded in other regions of the Iberian Peninsula: *Monfrague* National Park (Carranza *et al.* 1991) and *Sierra Morena* (Soriguer *et al.* 1994). The values of home ranges registered in these populations show that the northern part and the transboundary region of the study area gathered the habitat requirements for the species occurrence. The regions near the border with Spain have areas of contiguous habitat and low human pressure and could be considered a potential biological corridor for red deer dispersion enhancing the movement

among isolated habitat patches. The effective conservation of these areas can mitigate the pernicious effects of habitat fragmentation on population dispersion. This fact reinforces the importance and the need for an Iberian cross-border management of red deer populations emphasizing sustainable practices that enable the conservation and exploitation of this species. The correct population management coupled with a detailed knowledge about species dynamics is crucial to improve the accuracy of predictive models and allows the adoption of correct management approaches that reflects the balance between economic value and impact. Moreover, habitat suitability models are important tools, which can provide valuable information to achieve important management goals (e.g. trees protection for commercial forestry, conservation of the habitats and management for hunting exploitation (Pérez-Espona *et al.* 2009)). The success rate obtained in the model validation shows that the correct selection of the environmental variables is reflected in the quality and predictive power of the final output. Due to subjectivity associated with MCDM approach, the adoption of a Saaty's pairwise matrix (Saaty & Vargas 1991). and the final validation of the results are crucial to reduce the bias related to the predictions. Testing the same variables in a logistic regression approach we obtained an AUC value lower than for analytic hierarchic process (0.69 against 0.72, respectively) therefore we believe that the heuristic model applied in this study can be an alternative useful method. While AHP depends on the expert opinions (knowledge-driven), logistic regression is relied only on the data (data-driven). Logistic regressions, being a data-driven method where the the relative importance of the data is determined by the data itself should be the selected method when species data (presence, presence-absence) is available. The AHP method associated with GIS environment enhances the predictive performance of MCDM. We are aware that the model can be implemented in a larger scale in order to understand the species distribution and to identify possible biological corridors. Despite this, additional research is needed for testing other variables. Future investigation should be able to include predator activity as well as interspecific competition. Some variables that were used (*i.e.* land use) need to be updated to reduce uncertainties of the results. The variables selection and respective weight assignment were based in specialized bibliography and in our own field experience, however subjective judgment is a reality. The model validation was made using the data obtained by a direct count method. In such methods, the detection probability varies between open and closed areas. This fact, can introduce a possible bias that potentially reduces the effectiveness of the model evaluation. In order to reduce the uncertainties of direct

observations, future models should be validated through the application of indirect records of the species presence.

This study is a preliminary effort to understand and model the red deer habitat relationship in the southwestern edge of its European distribution. The methodology approach offers a simple and flexible way to meet the objectives set for this study. The habitat suitability maps produced in this study can be easily interpreted by non-experts (e.g. wildlife rangers, hunters) and landowners. In addition, as mentioned by Chen *et al.* (2010), the multi-criteria decision making using the analytical hierarchy process allows understanding how changes in the variable weights are reflected in the quantitative and spatial results.

Being aware of the uncertainties related with habitat suitability maps, we believe that expert knowledge coupled with GIS environment and field data (use to validate the final model) provides a quick and effective analysis of species-habitat relationships.

Acknowledgments

We are grateful to two anonymous referees for the review and helpful comments on the manuscript.

References

- Acevedo P. & Cassinello J. (2009). Human-induced range expansion of wild ungulates causes niche overlap between previously allopatric species: red deer and Iberian Ibex in mountainous regions of southern Spain. *Annales Zoologici Fennici*. 46: 39–50.
- Acevedo P., Farfán M.A., Márquez A.L., Delibes-Mateos, M., Real R. & Vargas J.M. (2011). Past, present and future of wild ungulates in relation to changes in land use. *Landscape Ecology*. 16: 19–31.
- Aguar C. (2000). *Flora e vegetação da serra de Nogueira e do Parque Natural de Montesinho*. PhD Thesis, Technical University of Lisbon. Portugal.
- Allen A.W., Jordan P.A. & Terrell J.W. (1987). Habitat Suitability Index Models: Moose, Lake Superior Region. Biological Report. U.S. Department of the Interior. Fish and Wildlife Service, Research and Development. Washington, DC. USA. 47 pp.
- Alvarez G. & Ramos J. (1991). Estrategias alimentarias del ciervo (*Cervus elaphus* L.) en Montes de Toledo. *Doñana, Acta Vertebrata*. 18: 63–99.
- Araújo M.B. & New M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*. 22: 42–47.
- Beaumont L.J., Hughes L. & Poulsen M. (2005). Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling*. 186: 250–269.
- Boitani L. (2000). Action Plan for the Conservation of Wolves (*Canis lupus*) in Europe. Nature and Environment, Council of Europe Publishing. 84 pp.
- Carranza J., Trucios S.J., Medina R., Valencia J. & Delgado J. (1991). Space use by red deer in a Mediterranean ecosystem as determined by radio-tracking. *Applied Animal Behaviour Science*. 30: 363–371.

Carvalho J. (2011). Monitorização da população de veado (*Cervus elaphus* L.) na Zona de Caça Nacional da Lombada, Nordeste Transmontano: Época 2010-2011. First degree thesis in Biology. University of Aveiro. Portugal.

Chen Y., Yu J. & Khan S. (2010). Spatial sensitivity of multi-criteria weights in GIS-based land suitability evaluation. *Environmental Modelling & Software*. 25: 1582–1591.

Clutton-Brock T.H., Guinness F.E. & Albon S.D. (1983). The costs of reproduction to red deer hinds. *Journal of Animal Ecology*. 52(2): 367–383.

Conolly J. & Lake M. (2006). *Geographical Information Systems in Archaeology*. Cambridge Manuals in Archaeology. Cambridge University Press, United Kingdom. 358 pp.

Debeljak M., Saso D., Klemen J., Andrej K. & Adamie M. (2001). Habitat suitability modelling for red deer (*Cervus elaphus* L.) in South-central Slovenia with classification trees. *Ecological Modelling*. 138(1-3): 321–330.

Dziba L.E., Scogings P.F., Gordon I.J. & Raats J.G. (2003). The feeding height preferences of two goat breeds fed *Grewia occidentalis* L. (Tiliaceae) in the Eastern Cape, South Africa. *Small Ruminant Research*. 47(1): 31–38.

Epperson B.K. (2000). Spatial and space-time correlations in ecological models. *Ecological Modelling*. 132: 63–76.

Ferreira S. (1998). Estudo da dieta de duas populações de veado (*Cervus elaphus* Linnaeus, 1758) em Portugal. First degree thesis in Biology, University of Coimbra, Portugal.

Fielding A.H. & Bell J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*. 24: 38–49.

Guisan A. & Thuiller W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*. 8: 993–1009.

Guisan A. & Zimmermann N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*. 135: 147–186.

Hewinson A.J.M., Vincent J.P., Joachim J., Angibault I.M., Cargnelutti B. & Cibiën C. (2001). The effects of woodland fragmentation and human activity on roe deer distribution in agricultural landscapes. *Canadian Journal of Zoology / Revue Canadienne de Zoologie*. 79 (4): 679–689.

Hirzel A.H., Le Lay G., Helfer V., Randin C. & Guisan A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*. 199: 142–152.

Jenks G.F. & Caspall F.C. (1971). Error on choroplethic maps: definition, measurement, reduction. *Annals of the Association of American Geographers*. 61: 217–244.

Jiang G., Ma J., Zhang M. & Stott P. (2009). Multiple spatial-scale resource selection function models in relation to human disturbance for moose in Northeastern China. *Ecological Research*. 24: 423–440.

Lima S.L. & Dill L.M. (1990). Behavioral decisions made under the risk of predation: A review and prospectus. *Canadian Journal of Zoology / Revue Canadienne de Zoologie*. 68: 619–640.

Malo J.E. & Suárez F. (1995). Herbivorous mammals as seed dispersers in a Mediterranean dehesa. *Oecologia*. 104: 246–255.

Marques F.F.C., Buckland S.T., Goffin D., Dixon C.E., Borchers D.L., Mayle B.A. & Peace A.J. (2001). Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland. *Journal of Applied Ecology*. 38: 349–363.

Moreira L.M., Rosa J.L., Lourenço J., Barroso I. & Pimenta V. (1997). Projecto Lobo. Relatório de Progressão 1996 (Cofinanciado pela U. E. - Programa LIFE). Parque Natural de Montesinho, Bragança, Portugal. 61 pp.

- Mysterud A., Langvatn R., Yoccoz N.G. & Stenseth N.C. (2002). Large-scale habitat variability, delayed density effects and red deer populations in Norway. *Journal of Animal Ecology*. 71: 569–580.
- Pearce J. & Ferrier S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*. 133: 225–245.
- Pérez-Espona S., Pemberton J.M. & Putman R. (2009). Red and sika deer in the British Isles, current management issues and management policy. *Mammalian Biology*. 74: 247–262.
- Phillips S.J., Anderson R.P. & Schapire R.E. (2006). Maximum entropy modelling of species geographic distributions. *Ecological Modelling*. 190: 231–259.
- Post E. & Forchhammer M.C. (2002). Synchronization of animal population dynamics by large-scale climate. *Nature*. 420: 168–171.
- Putman R. & Flueck W.T. (2011). Intraspecific variation in biology and ecology of deer: magnitude and causation. *Animal Production and Science*. 51(4): 277–291.
- Radeloff V.C., Pidgeon A.M. & Hostert P. (1999). Habitat and population modelling of roe deer using an interactive geographic information system. *Ecological Modelling*. 114: 287–304.
- Saaty T.L. & Vargas L.G. (1991). *Prediction, Projection and Forecasting*. Kluwer Academic Publisher, Dordrecht, Netherlands. 251 pp.
- San-José C., Braza F., Aragón S. & Delibes J.R. (1997). Habitat use by roe and red deer in southern Spain. *Miscellanea Zoologica*. 20: 27–38.
- Santos J. (2009). Estudo populacional do veado (*Cervus elaphus* L.) no Nordeste Transmontano. M.Sc. Thesis. University of Aveiro.

Schütz M., Risch A.C., Leuzinger E., Krüsi B.O. & Achermann G. (2003). Impact of herbivory by red deer (*Cervus elaphus* L.) on patterns and processes in subalpine grasslands in the Swiss National Park. *Forest Ecology and Management*. 181: 177–188.

Shi H., Laurent E.J., LeBouton J., Racevskis L., Hall K.R., Donovan M., Doepker R.V., Walters M.B., Lupi F. & Liu J. (2006). Local spatial modelling of white-tailed deer distribution. *Ecological Modelling*. 190: 171–189.

Soriguer R.C., Fandos P., Bernaldez E. & Delibes J.R. (1994). *El Ciervo en Andalucía*. Junta de Andalucía. Sevilla, Spain. 244 pp.

Stankovski V., Debeljak M., Bratko I. & Adamic M. (1998). Modelling the population dynamics of red deer (*Cervus elaphus* L.) with regard to forest development. *Ecological Modelling*. 108: 145–153.

Thapaliya K. (2008). Analysis of factors related to the distribution of red deer (*Cervus elaphus* L.) in Hustai National Park, Mongolia. M.Sc. Thesis. International Institute for Geo-Information Science and Earth Observation.

Thomas J.W., Black H., Scherzinger R.J. & Pedersen R.J. (1979). Deer and Elk. Pp. 104-127. In: J.W. Thomas (ed). *Wildlife habitats in managed forests of the Blue Mountains of Oregon and Washington*, Agricultural Handbook. Wildlife Management Institute, Washington, DC. USA.

Torres R.T., Virgós E., Santos J., Linnell J.D.C. & Fonseca, C. (2012). Habitat use by sympatric red and roe deer in a Mediterranean ecosystem. *Animal Biology*. 62: 351–366.

Vingada J., Fonseca C., Cancela J., Ferreira J. & Eira C. (2010). Ungulates and their management in Portugal. Pp. 392–418. In: M. Apollonio, R. Andersen, & R. Putman (eds). *European Ungulates and their Management in the 21st Century*. Cambridge University Press.

Capítulo 5

Tree-based methods in Species Distribution Models: How effective are they?

Carvalho J., Santos J., Acevedo P., Torres R.T., Gomes A. & Fonseca C. (*submitted*). Tree-based methods in Species Distribution Models: How effective are they?

Tree-based methods in Species Distribution Models: How effective are they?

Species distribution models (SDMs) have become widely used in the last decades and it has been recommended the simultaneous application of several methods (ensemble modelling). Concerning this, a plethora of techniques and algorithms has been developed in order to handle different aspects of conservation biology and wildlife management. Classification methods have become popular in recent years and their main advantages are (i) high classification accuracy, (ii) quantification and visualization of interactions between predictors, (iii) ability to handle non-linear relationships and unbalanced data and (iv) high predictive power. The aim of this study was to analyze the predictive performance of different modelling approaches prompting the use of alternative and promising methods for interpretation and prediction. Here, we explored the usefulness of ensemble modelling approaches. Four modelling methods were implemented to model the current distribution of red deer (*Cervus elaphus* L.) in an edge ecosystem, Portugal. Three statistical classifiers (classification and regression trees *CART*, boosted regression trees *BRT* and random forests *RF*) were applied and their predictive performance compared with a generalized statistical model. Firstly, a series of classification models were run with different tuning parameters trying to achieve the best performance. Secondly, models transferability among study areas was assessed. *RF* model was identified as the most accurate (AUC = 0.864; HL χ^2 = 13.8; p-value = 0.08) followed by *BRT* (AUC = 0.857; HL χ^2 = 11.8; p-value = 0.16), *GLM* (AUC = 0.742; HL χ^2 = 26.7; p-value = 0.001) and *CART* (AUC = 0.711; HL χ^2 = 44.7; p-value = $4.19e^{-07}$). Red deer occurrence and distribution was closely related with distance to perturbation areas, water bodies, shrubs, forest patches and topographic factors (e.g. slope). The results are discussed from an ecological and methodological perspective, highlighting the relevance of accurate predictions in the species management.

Keywords: *Cervus elaphus*, generalized linear model, machine learning, Portugal, species distribution models

Capítulo 5

Introduction

Understanding species distribution is a central topic in ecology. Species distribution models (SDMs) use environmental data and species records (e.g. species presence/absence) to predict other sites where the environmental conditions are suitable for their occurrence (see Elith & Leathwick 2009 and Zimmermann *et al.* 2010 for recent reviews). Such models are important tools in conservation planning and management (Schadt *et al.* 2002). Over the last decades, the enormous advancements in computing capacity and the accessibility to species occurrence data (e.g. atlas, large surveys, museum collections and online resources) have increased the number of modelling techniques available and, consequently, the number of publications on SDM approaches (Guisan & Zimmermann 2000; Elith *et al.* 2006; Austin 2007; Elith & Leathwick 2009). However, environmental data are complex and unbalanced requiring robust and discriminatory analytical methods to describe the interactions between species and environmental predictors. Nonetheless, in some cases, the conventional statistical methods are not flexible and the output results are difficult to interpret. Dealing with non-linear relationships, high dimensional data, heteroscedasticity and spatial autocorrelation are some of the major issues of ecological modelling (Olden *et al.* 2008). Despite the widely use of traditional methods e.g. generalized linear (McCullagh & Nelder 1989) and additive modelling (Hastie & Tibshirani 1990), the machine learning (henceforth ML) methods have become popular in recent years even though they are still outperformed by the previous ones. Contrastingly to the traditional statistical methods, ML uses an algorithm to record the relationship between ecological variables (Breiman 2001) describing dominant patterns and learning the relations between them and the response variable (Elith *et al.* 2008). Amongst the ML algorithms, the tree-based models are considered to have high capability to handle with interactions, nonlinearities and high data dimensionality, besides their high predictive power (see De'ath & Fabricius 2000; Prasad *et al.* 2006; De'ath 2007). Furthermore, these methods are non-parametric, do not require previous variable selection and are able to handle with missing values, outliers and unbalanced data. Notwithstanding all their advantages, ML exhibits some drawbacks, since they are time consuming and there is a lack of fundamental background to implement these methods.

The number of available techniques to model species distribution is increasing (Guisan & Thuiller 2005) and their performance has been evaluated (Elith *et al.* 2006; Tsoar *et al.* 2007). However, there are not conclusive results regarding the best method as the same methods have different performances with different datasets and in different regions (Elith *et al.* 2006). Following this line, it seems more appropriate to use various modelling methods to identify areas of consistent prediction (Araújo *et al.* 2006; Marmion *et al.* 2009).

Here, we use a red deer (*Cervus elaphus*) population as a case study to compare three different methods, based on a large database containing, approximately, six hundred records related with the species presence. Thus, the purpose of this study is to (i) assess the distribution of red deer in an edge ecosystem, considering three modelling approaches that rely on dichotomous events (e.g. presence-absence): Classification and Regression Trees (CART, Breiman *et al.* 1984; Clark & Pregibon 1992); Random Forests (RF, Breiman 2001) and Boosted Regression trees (BRT, Freund & Schapire 1997), (ii) select the most accurate SDM technique to model red deer distribution; (iii) assess the relative importance of the environmental variables for each data subset and (iv) evaluate the accuracy of the models to predict species distribution areas. Another aim was to compare these methods' performances with one of the most widely applied modelling technique, the Generalized Linear Model (GLM, McCullagh & Nelder 1989). These methods were developed in one region and then their spatial transferability was tested to another study area outside their original scope. To achieve this, a four-fold framework was used: i) delimitation of geographical background (GB) by applying a trend surface analysis (TSA) (Acevedo *et al.* 2012); ii) data selection and models development; iii) assessment of space-transferability of the GLM and tree-based approaches and iv) evaluation and selection in order to provide a statistical support to compare the modelling techniques applied. This will allow to understand: i) how effective tree-based methods are in predicting geographic patterns of occurrence; ii) the main drivers that affect the red deer occurrence (classification predictors) and iii) if tree-based models are transferable in space with reliability.

Material and methods

Study area

The study was carried out in two areas: i) Northwestern Iberian Peninsula where the models were developed (*Lombada-Sierra de la Culebra* hunting areas) and ii) *Lousã* National Hunting Area (NHA) (Center of Portugal) where the models were transferred (Figure 20).

The vegetation in Northwestern Iberian Peninsula, although dominated by scrubland, is varied and characterized by Pyrenean oak (*Quercus pyrenaica*), sweet chestnut (*Castanea sativa*), scots pine (*Pinus sylvestris*), European black pine (*Pinus nigra*), maritime pine (*Pinus pinaster*), and holm oak (*Quercus rotundifolia*). Main understory species are *Erica australis*, *Cystisus* sp., *Pterospartum tridentatum*, *Halimium lasianthum* subsp. *alyssoides* and *Cistus ladanifer*. Scattered pastures and agricultural crops can still be found along the study area. Two other wild ungulates are present: roe deer (*Capreolus capreolus*) and wild boar (*Sus scrofa*). Iberian wolf (*Canis lupus signatus*) is the main predator of red deer and densities have been calculated in 1.6 – 3.1 wolves/100 km² (Moreira *et al.* 1997). Human population density (9.5 inhabitants /km²) as anthropogenic disturbance is low.

According to Alves *et al.* (2013), *Lousã* NHA is characterized by three major habitat types: shrubland, coniferous forest and mixed forest. Shrubs are dominated by *Calluna vulgaris*, *Erica* spp., *Pterospartum tridentatum* and *Ulex* spp. The species *Pinus pinaster*, *Pinus sylvestris*, *Pinus nigra*, *Pseudotsuga menziesii* and *Cupressus lusitanica* are the most representative species of the coniferous forest. The mixed forests are represented by these conifer trees and some broadleaf species as *Quercus* sp., *Castanea sativa*, *Prunus lusitanica* and *Ilex aquifolium*. Monocultures of *Eucalyptus globulus* and exotic species as *Acacia* sp. are also common. Roe deer and wild boar are also present however there are not natural predators of these ungulates. Human population density (29.2 inhabitants/km²) is considerably higher than in *Lombada-Culebra* hunting areas.

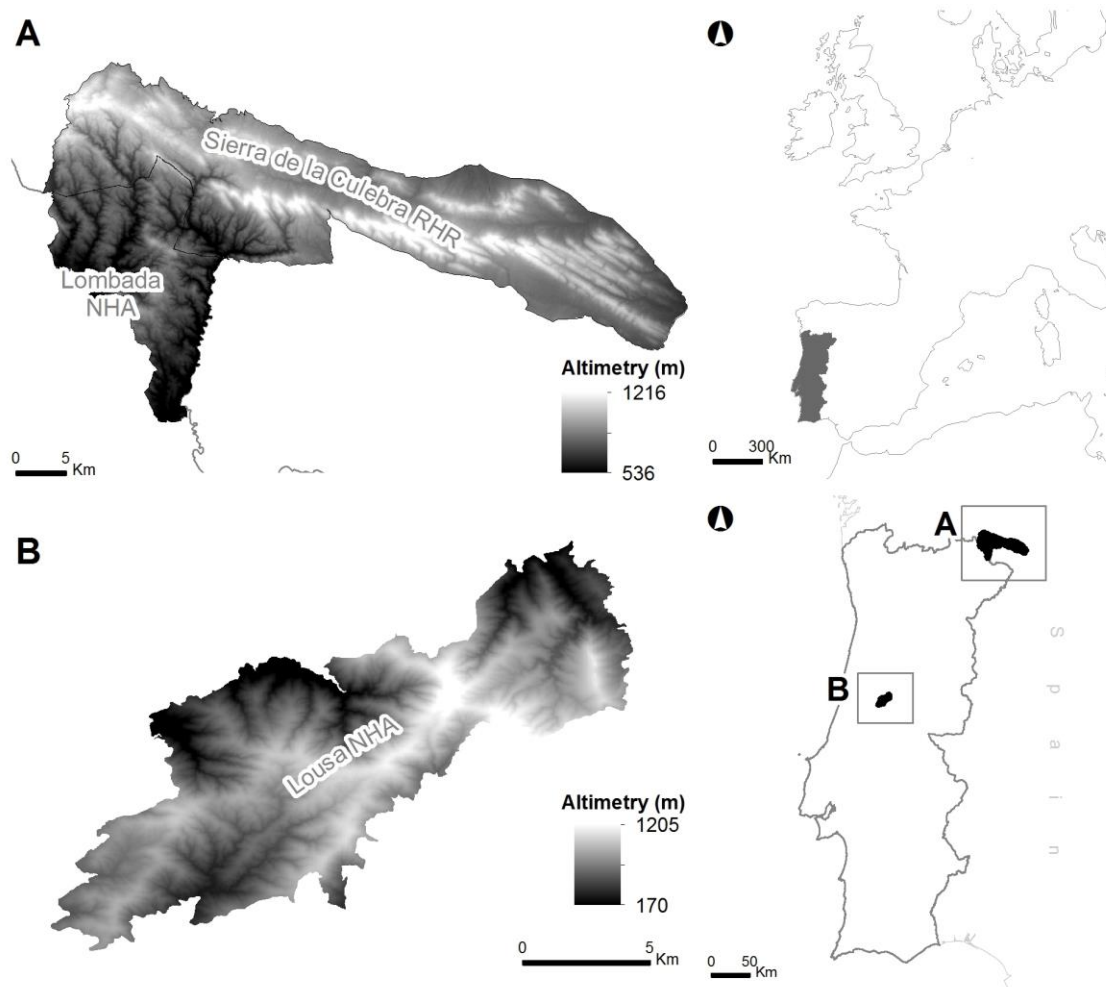


Figure 20 – Location of the study areas. A) *Lombada* National Hunting Area and the *Sierra de la Culebra* Regional Hunting Reserve where the models were developed and B) *Lousã* National Hunting Area where the models were projected.

Study species

Over the last decades, allied to the recent socioeconomic changes (e.g. rural exodus, abandonment of agricultural fields and renaturalization of the habitats) red deer populations have increased in number and distribution range in Portugal (Vingada *et al.* 2010). Since 1970's, these populations have remarkably recovered from a situation of virtual extinction being now common in many regions of Portugal, particularly along the Spanish border (Vingada *et al.* 2010).

Geographic background (GB) delimitation

Variations in the GB extent can be translated into differences in the discriminatory power of SDMs (Acevedo *et al.* 2012): if too restricted, GB can underestimate the role of coarse-scale factors (e.g. climatic drivers) (Jiménez-Valverde *et al.* 2011; Sánchez-Fernández *et al.* 2011), while if too large, GB can limit the model predictive power when the aim is to determine the influence of fine-scale conditions in geographic patterns of species distribution (Lobo *et al.* 2010). Here, we adopted the methodology proposed by Acevedo *et al.* (2012), using a geographical criterion in the GB delimitation. By applying a trend surface analysis (TSA) we aimed to improve the model performance and decrease the extent effect on the final outcome. TSA fits a polynomial surface by least-squares regression of geographic coordinates. This method is used to find general data tendencies and the modulation of curvilinear structures is performed through the addition of polynomial terms to the explanatory data (Legendre & Legendre 1998). Thus, trend surfaces are normally formulated as *n*th polynomials creating gradually varying surfaces that describe the physical or geographical processes. We fitted several surfaces increasing the polynomial order and, consequently, the complexity. The root mean square (RMS) error of interpolation was used to determine the best value to use for the polynomial order.

Data collection

In total, 48 line transects, with an average length of 5.7 km (2.4 – 12.7), were systematically distributed along the study area in order to provide an equal coverage of the most representative habitats. This survey design allows reduction of the bias associated with the systematic prospection of areas with different deer densities. When an animal/group was detected, the distance from the observation point to the animal/group

was recorded. Through GPS location and trigonometric operations the exact position of the observed animals was estimated. For modelling purposes we used 539 records of red deer presence divided into training (70%) and test (30%) datasets. Presence-only methods have been widely used in SDM as they provide more reliable information (Warton & Shepherd 2010). However, presence-only data disregards all the environmental conditions available for the species (Phillips *et al.* 2009). To overcome this problem, most modelling approaches require further information characterizing the range of environmental conditions for a given spatial context. Such information is provided by a random set of points from the study area. We adopt the background concept since it is based on fewer assumptions and there are a high number of methods able to handle with the overlap between presence and background data (Phillips & Elith 2011). Presence records were used to establish the favourable conditions for the species occurrence while background data characterizes the environmental domain where the survey was carried out. The ratio of presences/background data was 1:1 to avoid the model sensitivity to the prevalence of the species.

The study area was divided into a hexagonal grid composed by individual units (area = 16.24ha) that retains the environmental characteristics of the corresponding section. Hexagonal units are more suitable for a range of ecological applications than the commonly used rectangular grids (Clausnitzer *et al.* 2009) and allow a better representation of the spatial heterogeneity.

Exploratory analysis and variables selection

Variables selection is a key aspect in the performance of predictive models although its importance has been neglected (Williams *et al.* 2012). Prior to all analyses, we performed Spearman's rank correlation tests to check multicollinearity between predictors. Whenever the correlation coefficient exceeded 0.8, the variable with the lower biological significance was dropped (Jiménez-Valverde & Lobo 2006; Real *et al.* 2008). Based on red deer ecological requirements, their applicability to the study area and potential predictive power (Post & Forchhammer 2002), we selected 15 variables that can influence red deer occurrence and distribution: habitat structure (3 variables), human disturbance (2 variables), land use (3 variables), vegetation productivity (1 variable), topography (5 variables) and water availability (1 variable) (see Table 11).

Table 11 – Factors and variables used to model the red deer distribution in *Lombada-Sierra de la Culebra* and *Lousã* NHA. It is also presented the average values and variation by individual hexagonal unit.

<i>Factor</i>	<i>Variable</i>	<i>Code</i>	<i>Range in LNHA/SCRHR</i>	<i>Range in Lousã NHA</i>
Habitat structure	Agricultural area (ha)	AAgr	2.96 (0.00 – 16.24)	0.07 (0.00 – 6.65)
	Forest area (ha)	AFor	4.02 (0.00 – 16.24)	8.23 (0.00 – 16.24)
	Shrub area (ha)	ASch	7.16 (0.00 – 16.24)	7.85 (0.00 – 16.24)
Human disturbance	Distance to road network (m)	DRoad	950 (15 – 4785)	329 (26 – 1598)
	Distance to villages (m)	DUrb	1754 (0 – 5078)	1249 (85 – 3258)
	Distance to agricultural fields (m)	DAgr	465 (0 – 3257)	1430 (29 – 3578)
Land use	Distance to forest (m)	DFor	357 (0 – 2650)	60 (0 – 537)
	Distance to shrubs (m)	DSch	136 (0 – 1511)	71 (0 – 735)
	Normalized Difference Vegetation Index	NDVI	0.30 (-0.15 – 0.51)	0.42 (0.10 – 0.59)
Topography	Altitud (m)	Alt	876 (566 – 1550)	746 (232 – 1170)
	Slope (degrees)	Slp	10 (0 – 27)	21 (8 – 34)
	Northness	NNESS	-0.08 (-0.92 – 0.96)	0.05 (-0.90 – 0.92)
	Eastness	ENESS	-0.01 (-0.90 – 0.93)	-0.06 (-0.99 – 0.94)
	Terrain roughness	RNESS	155 (4 – 326)	116 (62 – 179)
Water availability	Distance to main water lines (m)	DRiv	1327 (55 – 5035)	760 (88 – 2578)

Modelling methods

Classification (categorical variable) trees, boosted regression trees and random forests were used to identify the variables most correlated with red deer distribution and to produce spatial predictions of habitat use across Northwestern Iberian Peninsula and Lousã mountains, as these models have advantages for species distribution modelling (see Leathwick *et al.* 2006; De'ath 2007; Elith *et al.* 2008). These methods were compared with a traditional statistical method, the generalized linear model (GLM).

CART

Classification (categorical variable) and regression (numerical variable) trees consist of binary recursive partition of the data distribution defined by the explanatory variables and based on the determination of impurity measures (e.g. Gini index) (Breiman *et al.* 1984). The variance resulting from the partition of the data is as homogeneous as possible with respect to the dependent variable. It is a non-parametric method that does not require assumptions regarding the data distribution. CART's are flexible considering the response variable, are able to incorporate complex interactions among variables and the final outcome is easily interpreted (De'Ath & Fabricius 2000). The development of a CART model was performed in four main steps: i) partition of the training dataset; ii) fitting a model to the data considering previous data partitions; iii) stop when the residuals of the model are approximately zero or the number of remaining observations is quite low and iv) pruning the tree to avoid overfit based on the complexity parameter (cp). A pruned tree allows the systematic representation of the data as simple as possible increasing the accurate predicts of unobserved data. For each tree, a series of 10-fold cross-validations were run and the most frequent occurring tree size was chosen using the 1-SE rule (De'Ath & Fabricius 2000). Classification trees were fitted using the *rpart* library developed by Therneau *et al.* (2013) and plotted using the *rpart.plot* library, an enhanced version developed by Milborrow (2012) for R software (version 2.15.3; R Development Core Team 2013).

BRT

BRT models were built using an algorithm that combines regression trees and boosting (Friedman *et al.* 2000). In BRT sequences of models are constructed and successive trees change the observation weights to give more relevance to those

instances incorrectly classified (De'Ath 2007). The aim of BRT is to improve the model performance by combining a large number of trees, being more reliable to calculate an average of several predictions instead to define a single and precise prediction (Elith *et al.* 2008). The model fit over different trees is improved by considering past fits. The contribution of the added trees is reflected by the re-estimation of fitted values for each observation.

BRT is optimized with two main parameters: learning rate (*lr*), which determines how quickly the algorithm adapts to a training dataset and the contribution of each tree to the growing model, and tree complexity (*tc*), which estimates interactions between predictor variables. The robustness to overfitting with respect to misclassification error is one of the main advantages of BRT, however the algorithm implementation is the most computationally taxing of the four methods described in this paper. The model was fitted with 10-fold cross-validations, testing different values for *lr* and *tc*. BRT were built using the *gbm* library developed by Ridgeway (2013) for R software (version 2.15.3; R Development Core Team 2013).

RF

In RF each tree is independently constructed using several bootstrap samples and new trees did not depend on earlier ones. RF fits many classification trees to a dataset, and then combines the predictions from all the trees. The algorithm begins with the selection of many bootstrap samples from the data. A classification tree is fit to each bootstrap sample, but at each node, only a small number of randomly selected variables (e.g. the square root of the number of variables) are available for the binary partitioning. The final model is performed by aggregating a set of predictions, the individual trees (for more details see Cutler *et al.* 2007). The RF meta-algorithm retains the variables that provide more information in the discrimination of items classes (Evans *et al.* 2011). The idea of variable importance is: if the values of a variable are removed and the model accuracy doesn't change significantly, then the variable effect is negligible. RF is able to measure the contribution of predictors even if their effects are overshadowed by correlative relations (Strobl *et al.* 2009). The method performance depends on two fundamental parameters, i) the overall number of trees, which should increase as the number of variables/instances increase, and ii) the number of predictors considered for each split whose value resulted from the square root of the number of variables. The main advantages of RF are the variance reduction which provides stable predictions, the robustness to noise data (e.g. outliers) and the high predictive power (Cutler *et al.* 2007).

RF models were fitted using the square root of the number of variables and a series of 10 ensemble sizes (5, 10, 20, 50, 100, 200, 500, 1000, 2000 and 5000). The random forest library implemented in R by Liaw & Wiener (2012) was used in models development.

GLM stepwise

GLM can be described as a flexible generalization of the ordinary linear regression allowing non-normal distribution of response variable (McCullagh & Nelder 1989). Here, GLM was performed using stepwise variable selection which is a common implemented approach in distribution modelling (e.g. Araújo *et al.* 2005a; Barbosa *et al.* 2008). In this procedure the variables with additional contribution to the model are added at each step. This method is a modification of the forward selection since effects already in the model do not necessarily remain there. If at each step of stepwise, any effect in the model is negligible, then the least significant of these effects is removed. Correlation between variables affects the coefficients but not the model predictive performance. The most parsimonious model was selected based on Akaike's Information Criterion (Akaike 1974).

Models calibration and discrimination

Discrimination capacity is the ability of a model to distinguish between positive instances (e.g. presence of the species) and negative instances (e.g. absence of the species) (Jiménez-Valverde *et al.* 2013). In SDMs this measure is usually the only characteristic assessed, however calibration is another parameter that provides useful information about model performance.

Here, we used the area under the Receiver Operating Characteristic (ROC) curve (AUC) and calibration plots to measure discrimination power and quantify the agreement between estimated probabilities of occurrence and observed proportions of presences. Spearman's rank tests were used to assess the correlation between variables importance estimated by the models applied. AUC provides a way to compare classifiers testing the model accuracy and allowing their validation independent of distortions and potential bias (Fielding & Bell 1997). The output values vary between 0.5 (scores of the two groups do not differ) and 1 (non-overlapping of the group scores). The data were randomly partitioned for model evaluation: 70% of the records used as training data to construct the model and the remaining 30% set aside for testing to ensure that the model is not overfitted and is generalizable.

Transferability assessment

SDMs are usually evaluated within the region in which they were fitted and their applicability to other parts of a species' geographical range has rarely been assessed (although see Thomas & Bovee 1993; Randin *et al.* 2006; Barbosa *et al.* 2008; Cianfrani *et al.* in press). This is mainly due to environmental differences between the area where the model was developed and the area where the model will be projected (for more details see Randin *et al.* 2006). Notwithstanding, if both areas are environmentally similar models can be reliably transferred outside their development domain (Bulluck *et al.* 2006).

In this study, we tested the spatial transferability of the four models developed (*CART*, *BRT*, *RF* and *GLM*) using presence data of the projection area. The process was carried out in four main steps: i) we dealt with multicollinearity by dropping from the model all the predictor variables whose variance inflation factor (VIF) was higher than 5; ii) we evaluated the similarity between the two study areas through the application of *Mahalanobis* distance (Cianfrani *et al.* in press); iii) we assessed the correlation between the two environmental matrices using the Mantel test and iv) we evaluated the model predictive performance.

Results

Geographical background

Following the methodological approach suggested by Acevedo *et al.* (2012), a TSA using known occurrences of the species was performed, in order to reduce the extent effects on the model predictive performance. A root mean square error of 0.43 was obtained through the application of a third-order polynomial (Figure 21) which is the recommended for processes that occur at the same or higher range than the study region (Legendre & Legendre 1998). This area represents the geographical delimitation where the models was fitted and validated.

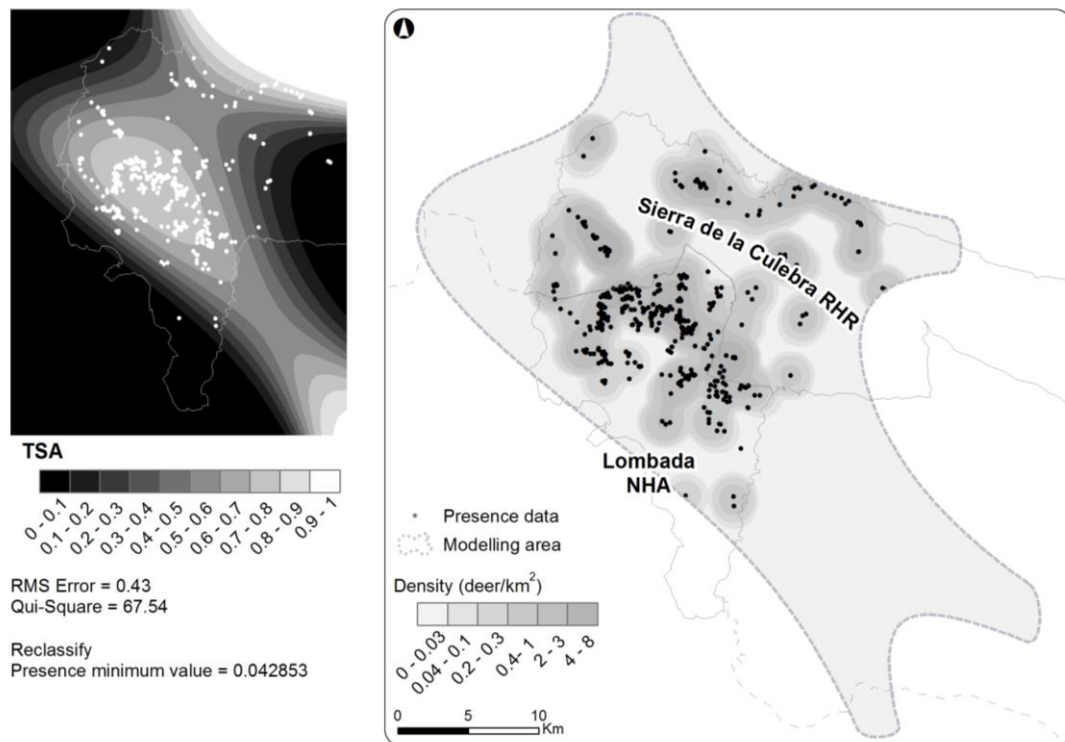


Figure 21 – Trend surface analysis using a third-degree polynomial for red deer presences.

CART

Classification tree (CT) models were developed with all the variables except *AFor* due to high collinearity with the variable representing the distance to forest patches ($r = 0.91$) (see correlation matrix in the appendix I). For the whole tree, 48 out of 135 instances were misclassified, giving an overall error of 36%. The first tree, constructed to its maximum depth, was pruned using the complexity parameter ($cp = 0.021$) avoiding the overfit of the training data and increasing the model performance when dealing with new records (Figure 22).

The cross-validation (CV) error was calculated using approximately two-thirds off all data reaching a minimum value of 0.61 for the largest tree of size 17. There is a relative reduction in the error as the size of tree increases and the complexity parameter decreases (Figure 22c). Based on the CV error and on the cp , the optimal tree size comprises six data splits and retains four variables: *DUrb*, *DSch*, *NNESS* and *NDVI*. Considering all the CT models developed these four variables were most frequently selected due to the amount of explained deviance. The relative high number of splits shows that the estimated response does not depend only on main effects.

Red deer absence is associated with the proximity to urban areas (below 1620 meters) (Figure 22d). When the distance is higher than 1620 meters, the distance to shrubs is the second most decisive variable followed by the degree of exposure to north and *NDVI*. Red deer presence is predicted by high distance to urban areas, low distance to shrubs and a negative exposure to north.

Applying regression tree models to predict red deer abundance (Figure 23), our results show that the most important variables are the distance to agricultural fields, forests and roads. High densities are observed far from agricultural fields while low distances to the forests patches are related with low densities. When distance to forests and distance to roads increased an increment in red deer densities was observed.

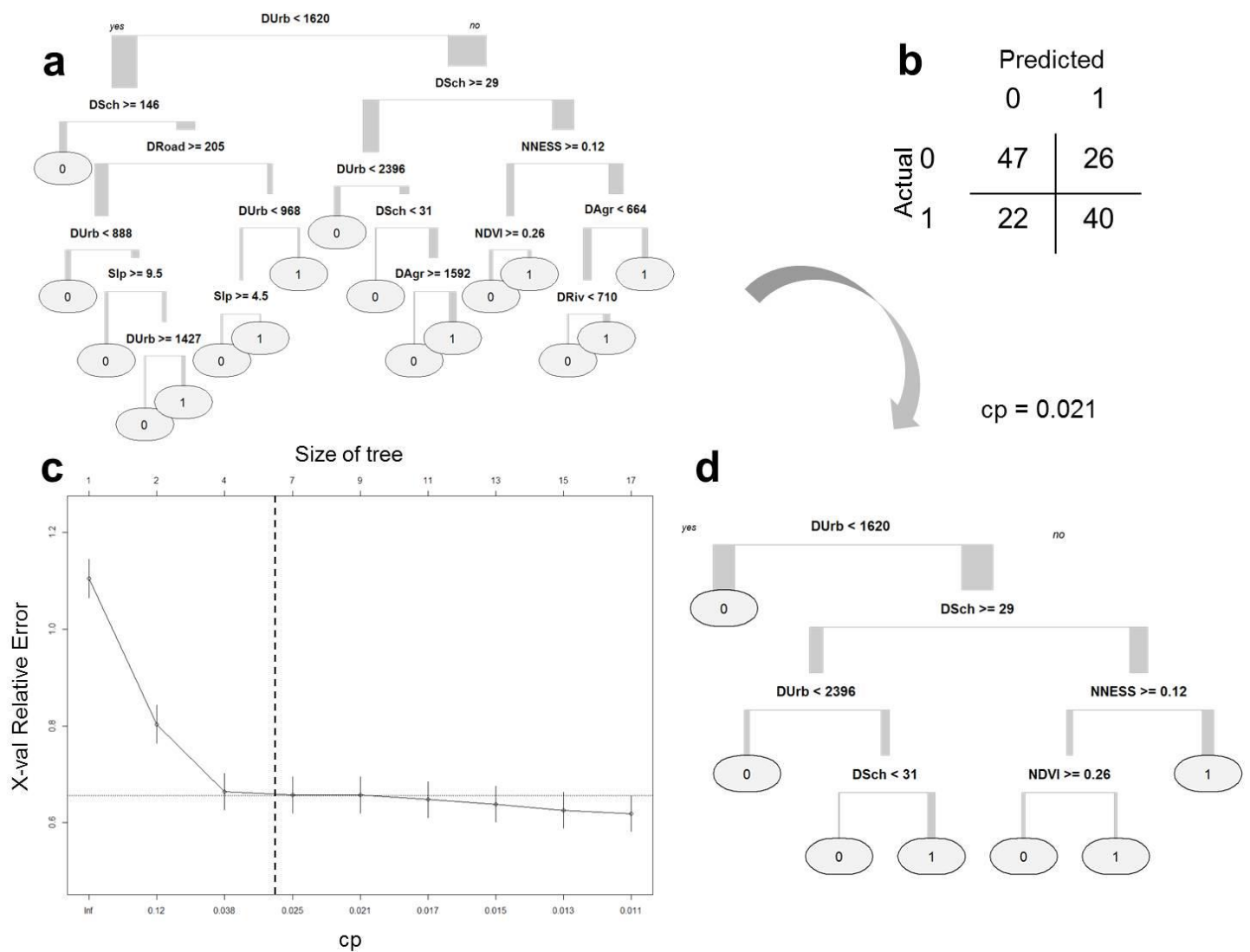


Figure 22 – The decision trees (a and d) are read from the top to bottom. The nodes represent the successive splits of the training dataset defined in order to increase the homogeneity between groups. The branches indicate the path taken by individual cases and the branch width provides a visual proportion of instances that were clustered on each side. Confusion matrix is represented in figure b and the figure c illustrates the relative reduction of cross-validation error in relation to the size of tree and complexity parameter.

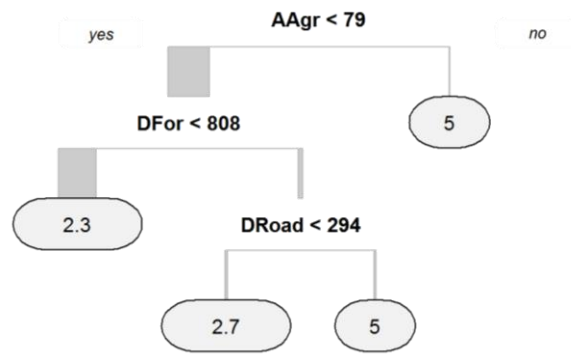


Figure 23 – Regression tree fitted to predict red deer abundance across the study area.

AUC for CT model is 0.711. The model calibration was assessed through calibration plot (Figure 23). The Hosmer-Lemeshow (HL) showed that the model is not perfectly calibrated showing significant deviations between predictive and observed values (HL: $\chi^2 = 44.7$ and $p\text{-value} = 4.2e^{-07}$).

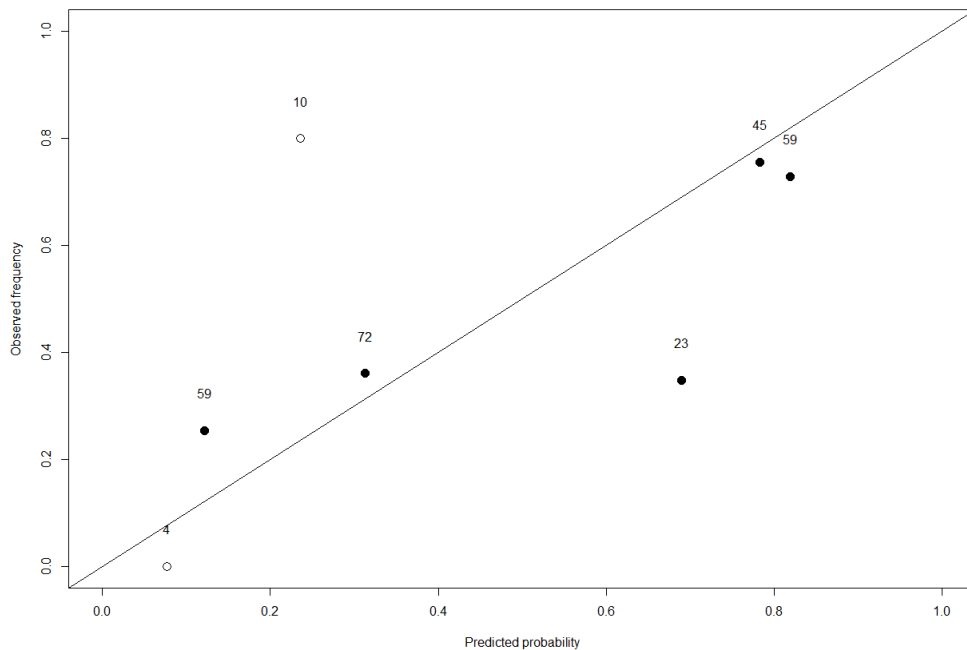


Figure 24 – Calibration plot showing the adjustment between predictive/estimated probabilities (x-axis) and the observed proportion of positive cases (y-axis).

BRT

The best BRT models were achieved using the follow parameters: learning rate (lr) of 0.005 and 0.01, a tree complexity (tc) of 5 reserving a bag fraction (bf) of 0.5. The model was built with 10-fold cross-validation with all the predictors included. The variable importance plot (Figure 25a) shows strong effects for distance to urban areas (*DUrb*), distance to rivers (*DRiv*), distance to shrubs (*DSch*), distance to roads (*DRoad*) and distance to agricultural fields (*DAgr*) which are partially in agreement with CT model. Moderate effects were observed for terrain roughness (*RNESS*), agricultural area within a buffer of 250 meters surrounding the presence records (*AAgr*), slope (*Slp*) and northness (*NNESS*). Finally, weak effects were observed for distance to forest patches (*DFor*), NDVI, eastness (*ENESS*) and for forest area within a buffer with the same characteristics as the mentioned above (*AFor*). Considering a lr of 0.01 (Figure 25b) the results are quite similar varying in the degree of importance assigned to only three variables (*DRoad*, *DAgr* and *DSch*).

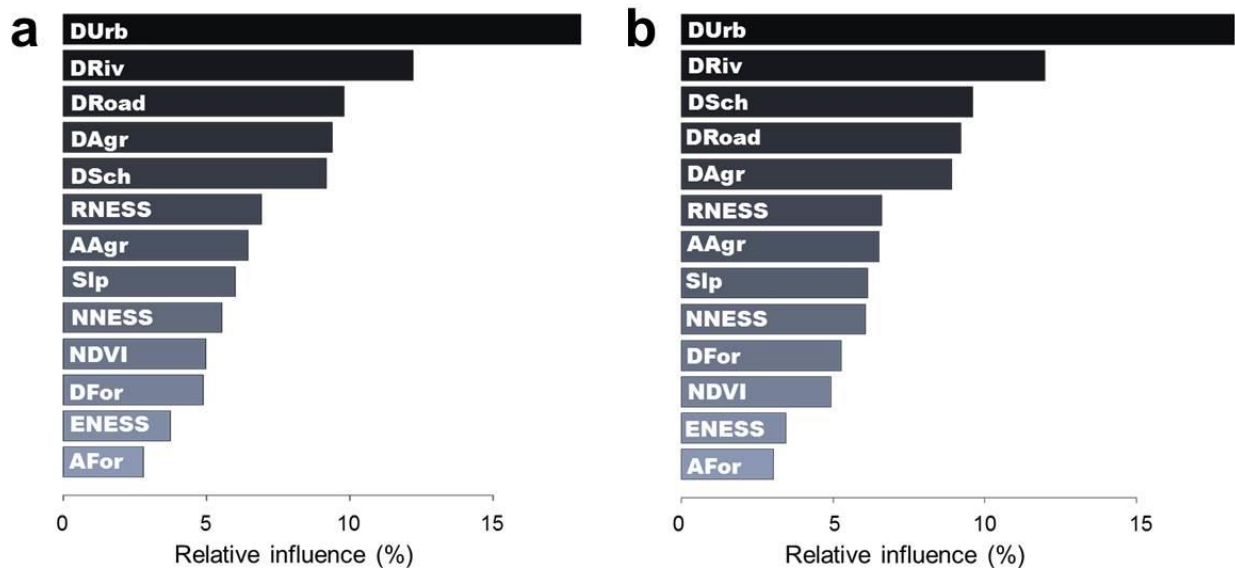


Figure 25 – Variable importance plot for a learning rate of 0.005 (a) and 0.01 (b) showing that the distance to urban areas is the most relevant variable affecting red deer occurrence.

The partial dependency plots showed: a predominantly linear positive trend and then a plateau for *DUrb*; downward trend followed by a plateau for *DSch*, *Slp* and *DFor*; weak effects for *RNESS*, *NNESS* and *ENESS*; and modal effects for the remaining variables (see Figure 26).

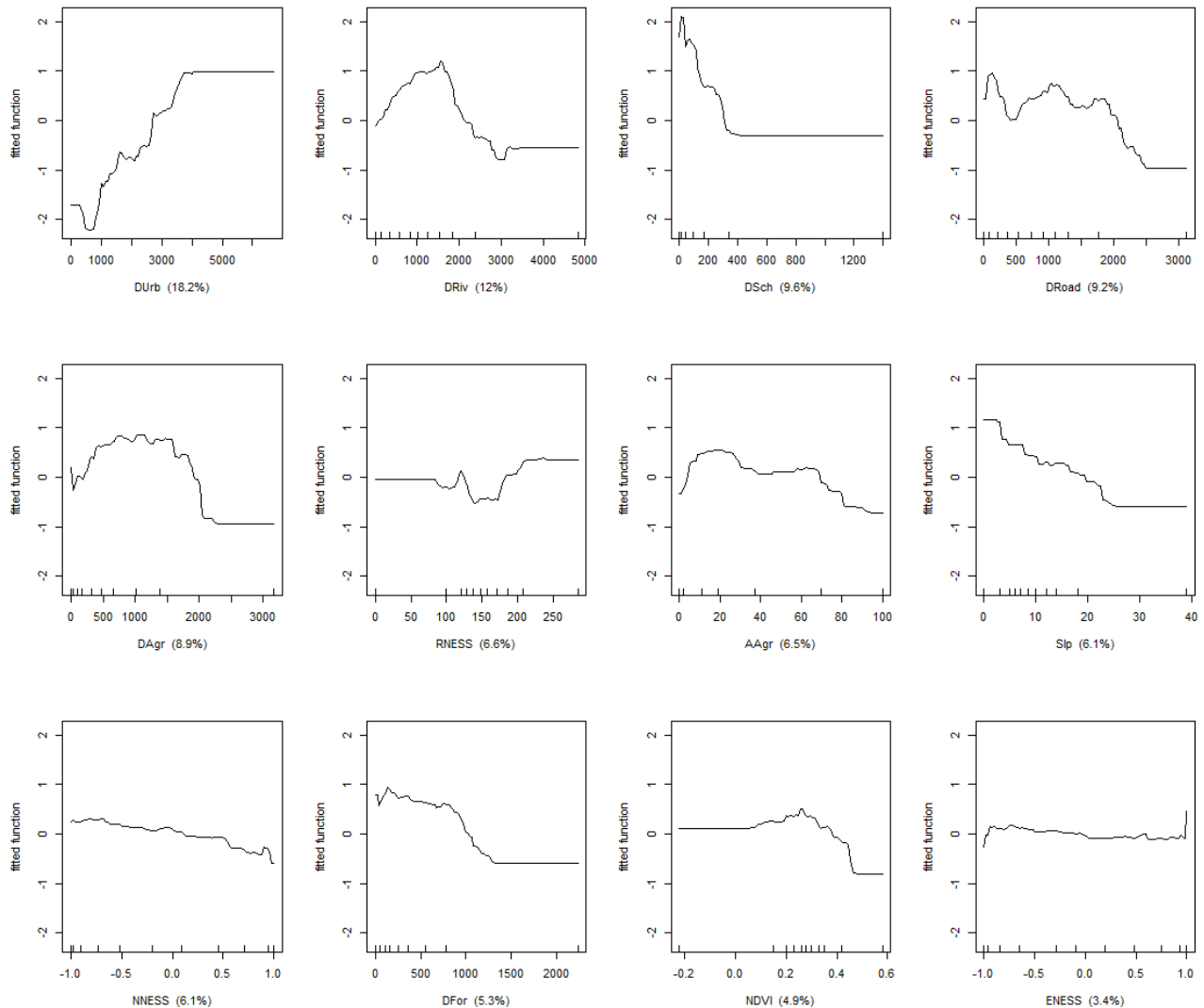


Figure 26 – Partial dependence plots for the twelve most relevant variables identified in BRT model affecting red deer occurrence. Variables explanation and range for the modelling area (*Lombada-Culebra* hunting areas) are presented in the table 11. Rug plots at bottom of plots show distribution of records across the variable range, in deciles.

The response of red deer for the 12 most influential variables indicates that the species occurs far from perturbation areas, near to water lines, shrubs and forest patches. The species often occurs in areas with smooth slopes. The exposure effects represented by *NNESS* and *ENESS* seem to have a little effect, although it is possible to describe a preference by hillsides exposed to south (Figure 27).

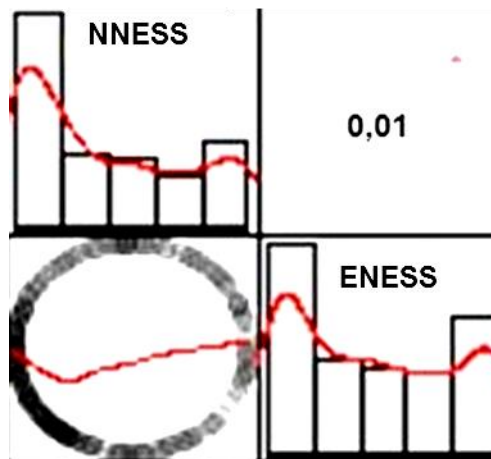


Figure 27 – Circular plot correlating the variables related with exposure effects. It is possible to observe an increase of presence records in the hillsides exposed to south and southwest.

AUC for BRT model was 0.857, registering a high improvement in relation to CT model. The calibration plot and HL values revealed a good adjustment between predicted and observed values (Figure 28).

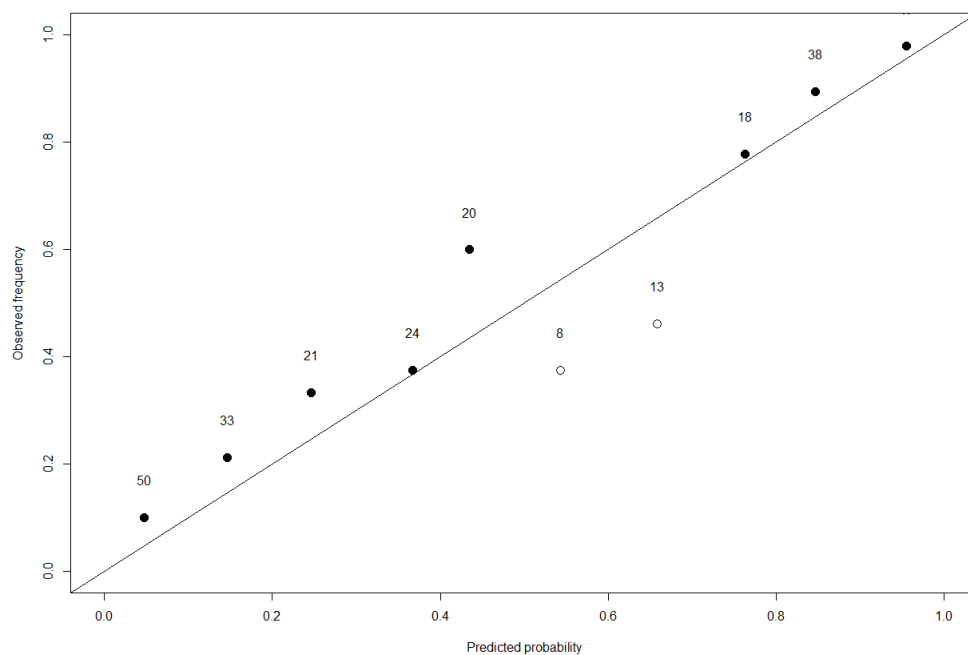


Figure 28 – Calibration plot showing the adjustment between predictive/estimated probabilities (x-axis) and the observed proportion of positive cases (y-axis). HL ($\chi^2=11.8$ and p-value=0.16) values revealed a good adjustment between predicted and observed values.

RF

The outcomes of RF model depend primarily on two parameters: the number of predictors chosen randomly at each fitted tree node and the overall number of trees. The results were achieved using one-third of all predictors per node and a variable number of trees (5, 10, 20, 50, 100, 200, 500, 1000, 2000, and 5000) in order to stabilize the error predictions. If there are noisy variables, a high number of trees is recommended (Figure 29). Figure 30 represents cartographically the predictive results. Increasing the number of trees the error decreases at a certain value as the AUC increases reaching a plateau at 200–500 trees (Out-of-bag (OOB) error=21.20 and AUC=0.858).

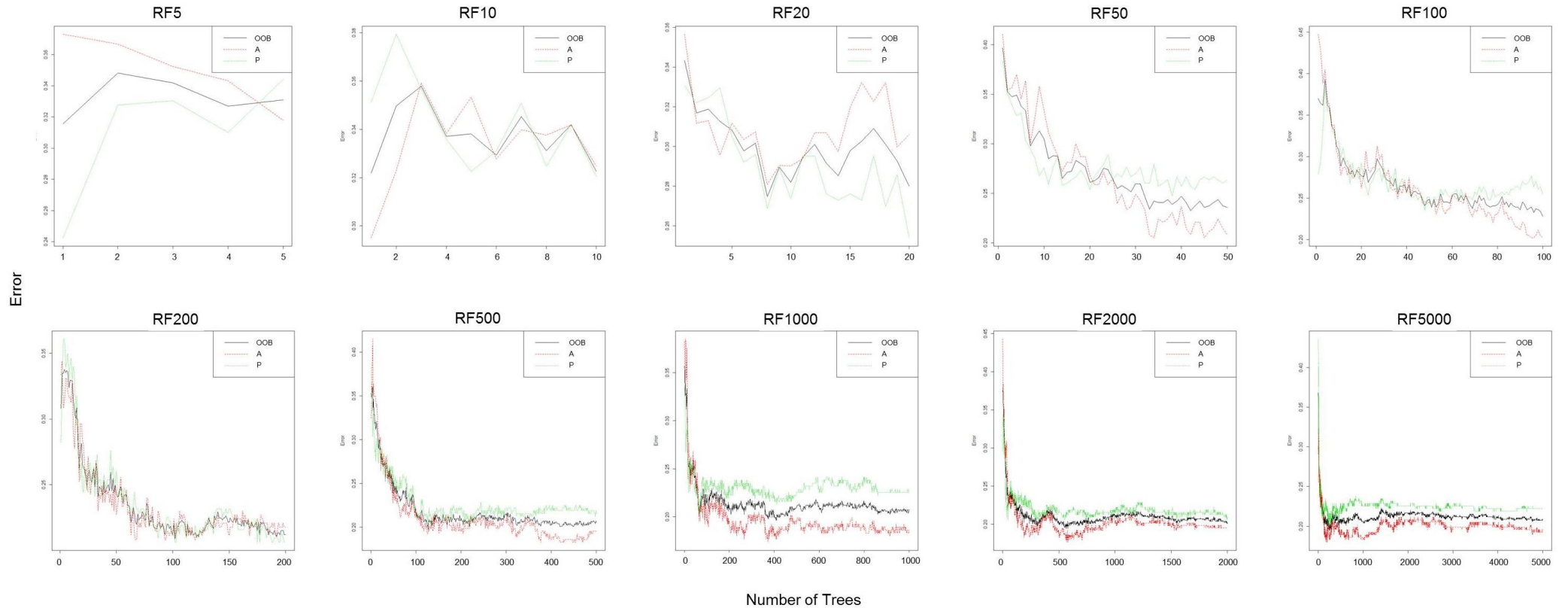


Figure 29 – The error rate of the overall model decreases as each new fitted tree is added to the ensemble. Reaching 200 trees the error rate does not decrease which indicates that the trees addition does not improve the model performance. OOB – Out-of-bag error; A – Error related with the absences represented by background data; P – Error related to the presence records.

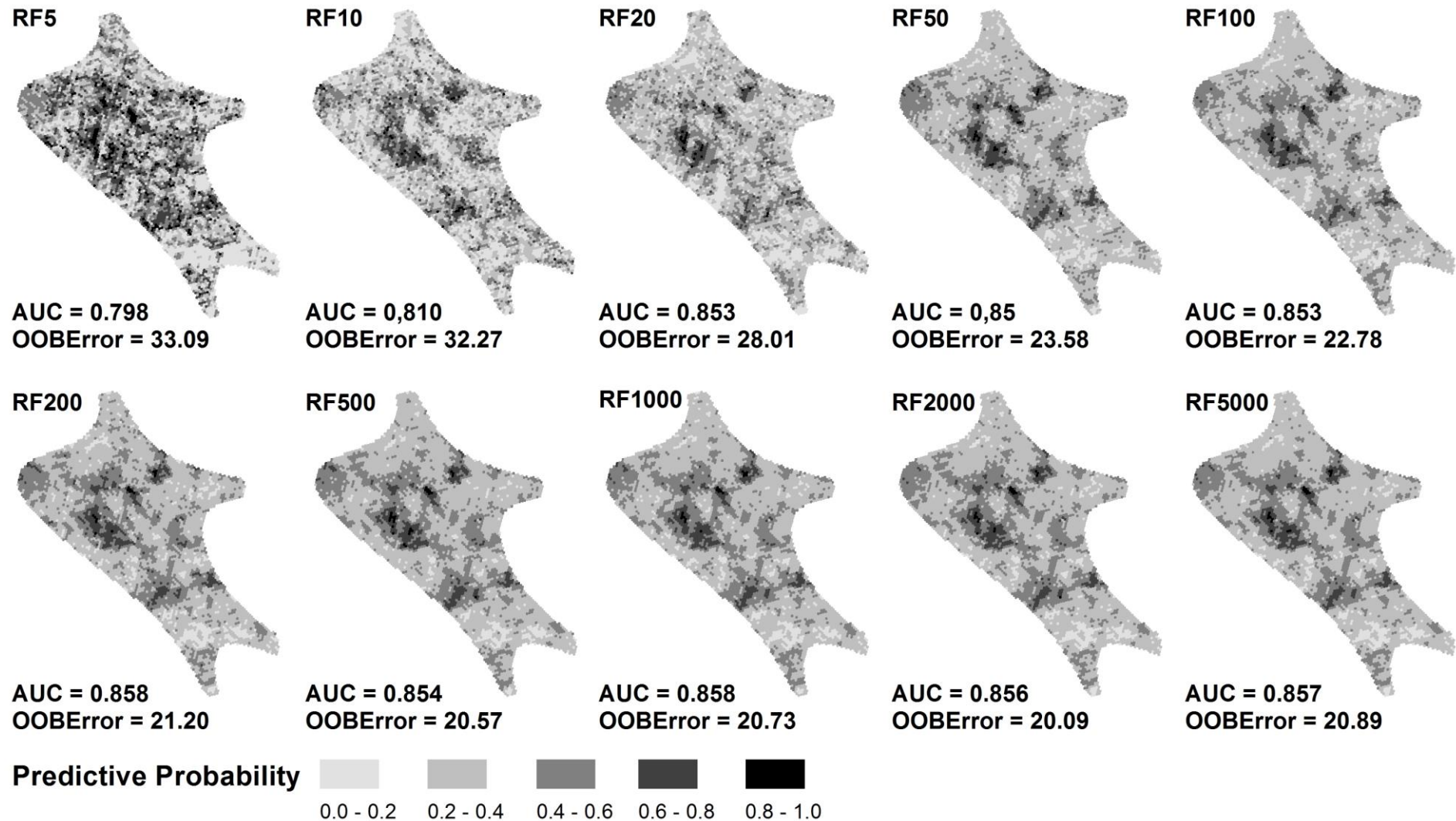


Figure 30 – Cartographic representation of random forest models. The model with 5 trees represents a diffuse pattern in the occurrence probabilities, however, increasing the number of trees added to the ensemble makes this pattern more stable and defined.

Figure 31 represents the relation between the OOB error and the discriminatory model performance.

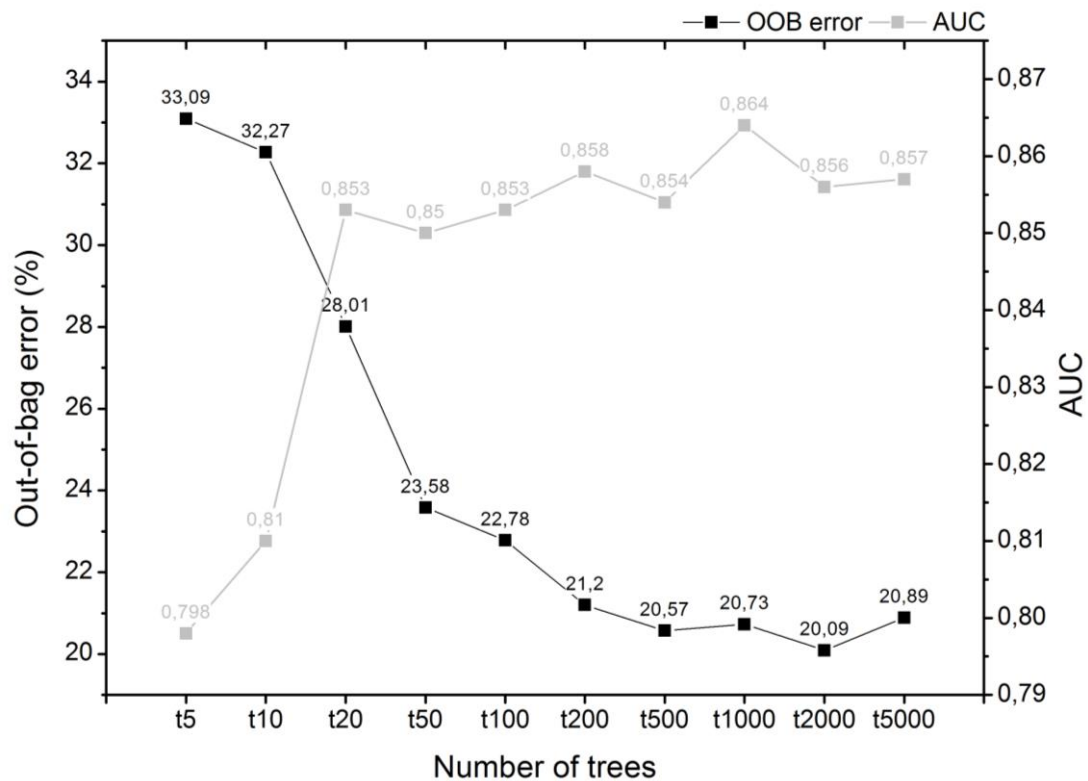


Figure 31 – Graphical representation of the OOB error and AUC values. Adding trees to the ensemble reduces the error and increases discriminatory model performance reaching a plateau at 200 trees.

Variable importance varies with the number of trees (Figure 32). However, as the number increases and the OOB error decreases the variables selected and their relative importance are close to the previous methods tested. Considering the model that fits 1000 trees (highest AUC and one of lowest OOB error), the variable with strong effects was the *DUrb*, followed by *DAgr*, *AAgr*, *DRiv*, *DSch* and *ASch*. The variables *DRoad*, *Slp*, *NDVI*, *DFor* and *AFor* were associated with moderate effects while variables related with terrain roughness (*RNESS*) and exposure (*NNESS* and *ENESS*) showed weak effects in the prediction of red deer occurrence.

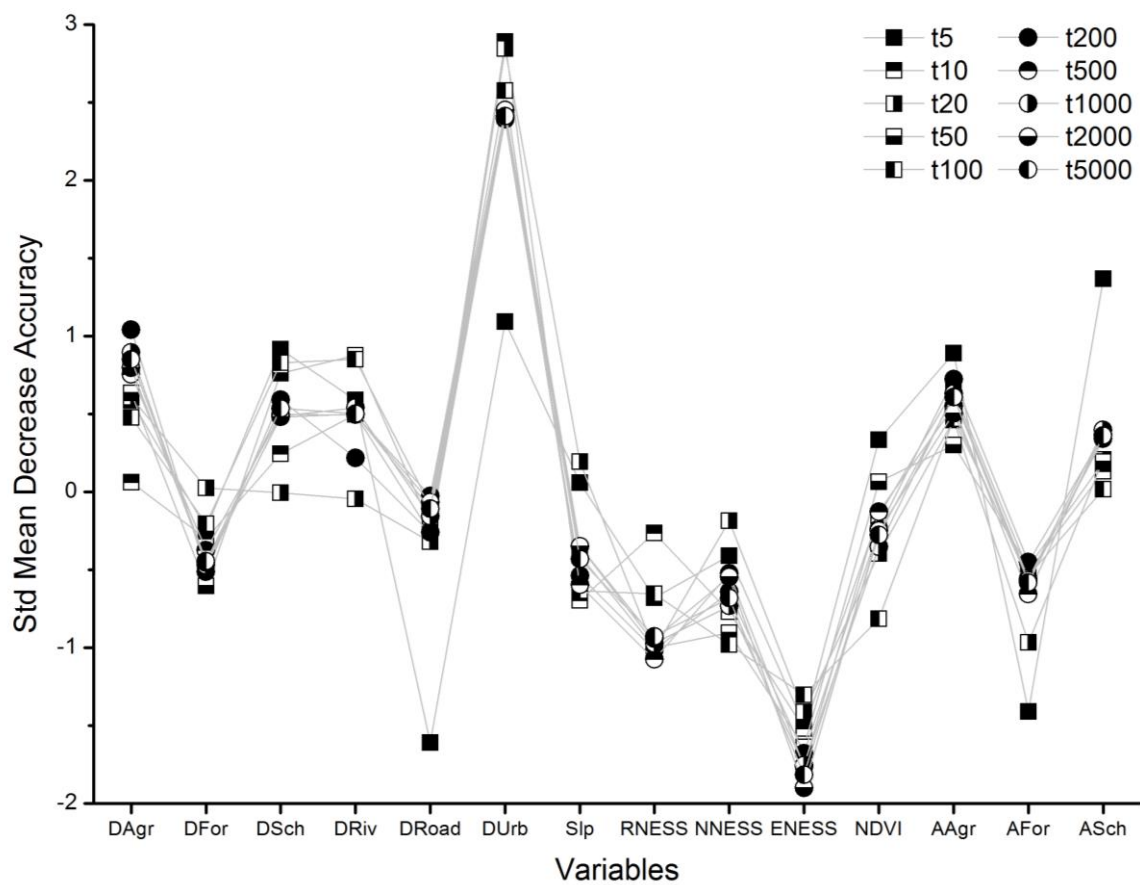


Figure 32 – Variable importance in relation to the number of fitted trees.

AUC for RF was 0.864, which reflects a slight increase in the predictive performance when compared to BRT model. The calibration plot and HL values revealed a good adjustment between predicted and observed values (Figure 33).

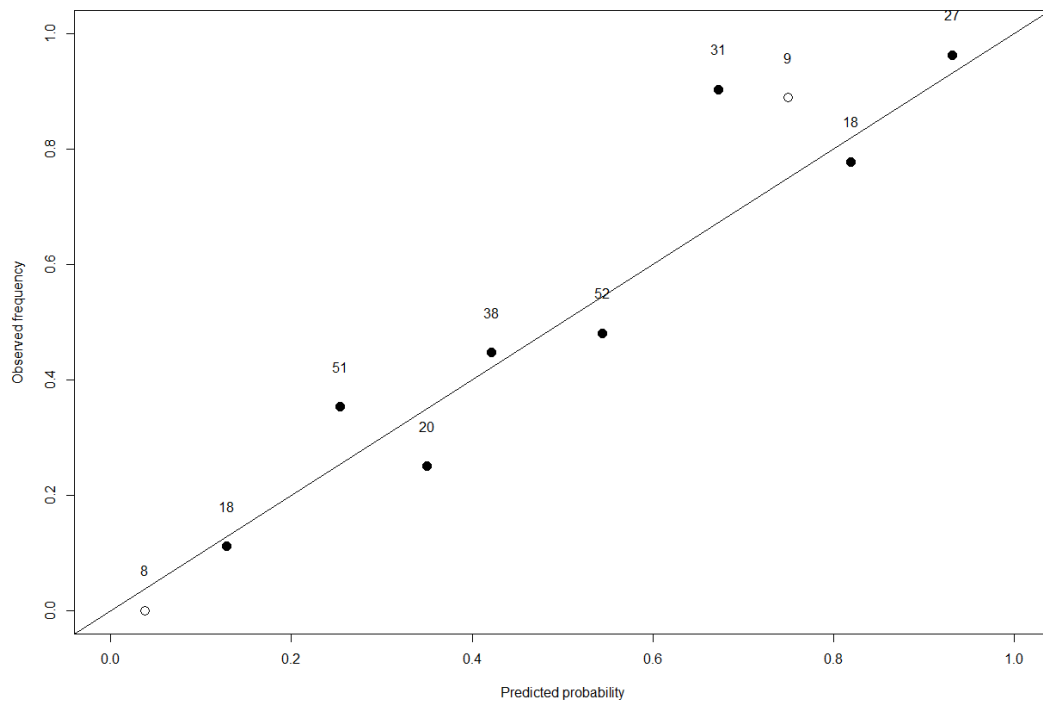


Figure 33 – Calibration plot showing the adjustment between predictive/estimated probabilities (x-axis) and the observed proportion of positive cases (y-axis). HL ($\chi^2=13.9$ and $p\text{-value}=0.08$) values revealed a good adjustment between predicted and observed values.

GLM Stepwise

The most parsimonious GLM stepwise model has retained 10 variables (Table 12). From the initial set of 15 environmental variables, the following factors were dropped: *DAgr*, *ASch*, *AFor*, *ENESS* and *Alt*. The model outcomes were partially in agreement with the remaining techniques applied. The *DUrb* had a positive effect in the occurrence probability of red deer (Table 12). For the *DSch*, *Slp*, *DRoad*, *AAgr*, *DRiv*, *NNESS*, *NDVI*, *DFor* and *RNESS* the effects were negative (Table 12). An increase in the measure units of these variables decreases the species occurrence probability.

Table 12 – Coefficients (β), standard deviation (*SD*) and the significance (*p*) for the regression model applied. Significance codes: “***” significant at 0.001, “**” significant at 0.01, “*” significant at 0.05, “.” significant at 0.1.

	<i>B</i>	<i>S.D.</i>	<i>z value</i>	<i>P</i>
<i>Constant</i>	0.6732	0.5001	1.346	0.1783
<i>DUrb</i>	0.0084	0.0001	7.515	5.69e-14 ***
<i>DSch</i>	-0.0036	0.0007	-5.511	3.56e-08 ***
<i>Slp</i>	-0.0846	0.0210	-4.038	5.39e-05 ***
<i>DRoad</i>	-0.0006	0.0002	-3.702	0.0002 ***
<i>AAgr</i>	-0.0087	0.0038	-2.292	0.0219 *
<i>DRiv</i>	-0.0003	0.0001	-2.597	0.0094 **
<i>NNESS</i>	-0.3134	0.1357	-2.310	0.0209 *
<i>NDVI</i>	-1.9960	0.9586	-2.082	0.0373 *
<i>DFor</i>	-0.0005	0.0003	-1.806	0.0710 .
<i>RNESS</i>	0.0044	0.0029	1.538	0.1239

The AUC for GLM stepwise was 0.742. This value, despite lower than ensemble models, represents an increase in the discrimination performance in relation to CT. The calibration plot and HL values revealed a slight improvement in the adjustment between predicted and observed values when compared to CT model.

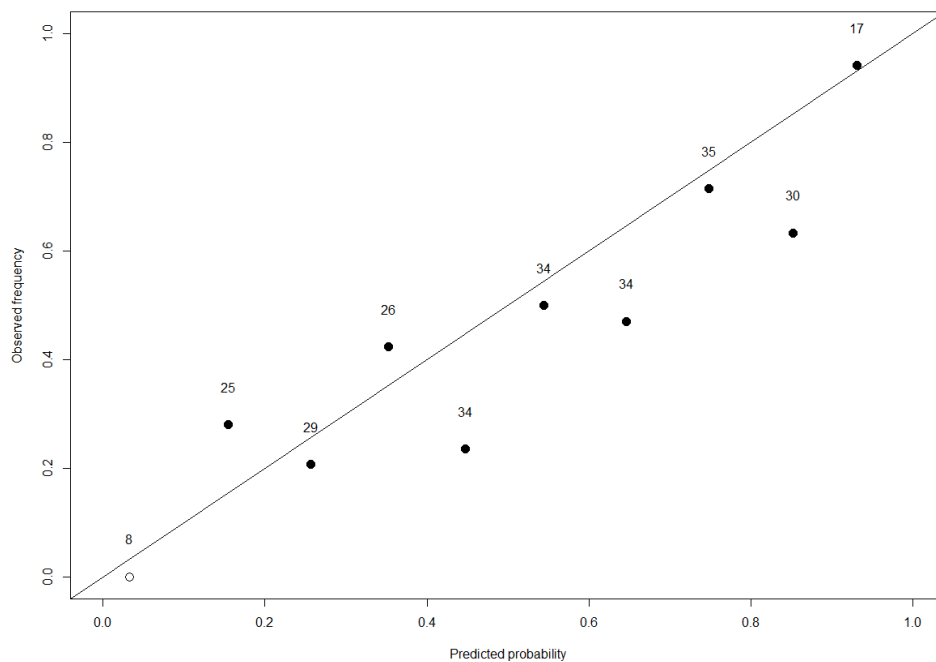


Figure 34 – Calibration plot showing the adjustment between predictive/estimated probabilities (x-axis) and the observed proportion of positive cases (y-axis). HL ($\chi^2=26.7$ and $p\text{-value}=0.001$).

To run the Spearman's correlation test, the variables were ranked according to their order of selection by each model. The results showed a positive correlation between models indicating a partial concordance of variable importance among the techniques applied (Table 13). Considering the results of all the models, variables were ranked in the following: *Durb*, *DSch*, *DRoad*, *DRiv*, *DAgr*, *Slp*, *AAgr*, *NNESS*, *NDVI*, *RNESS*, *DFor*, *ASch*, *AFor* and *ENESS*. The lowest correlation values were associated with RF model due to some considerable differences related with the order of variables selection. This method highlights the variables *AAgr* and *ASch* neglecting the *DRoad*, *NNESS* and *Slp* which represents a mismatch in relation to the others techniques.

Table 13 – Correlation matrix comparing the rank order of variable selection by the four modelling techniques applied.

	<i>CART</i>	<i>BRT</i>	<i>RF</i>	<i>GLM</i>
CART	1.00	0.73	0.50	0.69
BRT	-	1.00	0.62	0.65
RF	-	-	1.00	0.46
GLM	-	-	-	1.00

Transferability

Variance inflation factors applied to the *Lombada-Culebra* environmental matrix revealed multicollinearity issues among variables. Further, the Mahalanobis distance showed some dissimilarity between *Lombada-Culebra* and *Lousã* NHA. Thus, a total of six variables were dropped (*AFor*, *AAgr*, *RNESS*, *Slp*, *DSch*, *Alt*) in order to guarantee environmental similarity between the study areas. The Mantel test showed that differences between matrices were not statistically significant (p-value=0.07), which supports the transferability of the models from *Lombada-Culebra* to *Lousã* NHA. Models developed with different methodological approaches can be considered potentially useful since they have AUC values higher than 0.7. Following the projection in *Lousã* NHA, all the models were validated with records related to species presence (Figure 35).

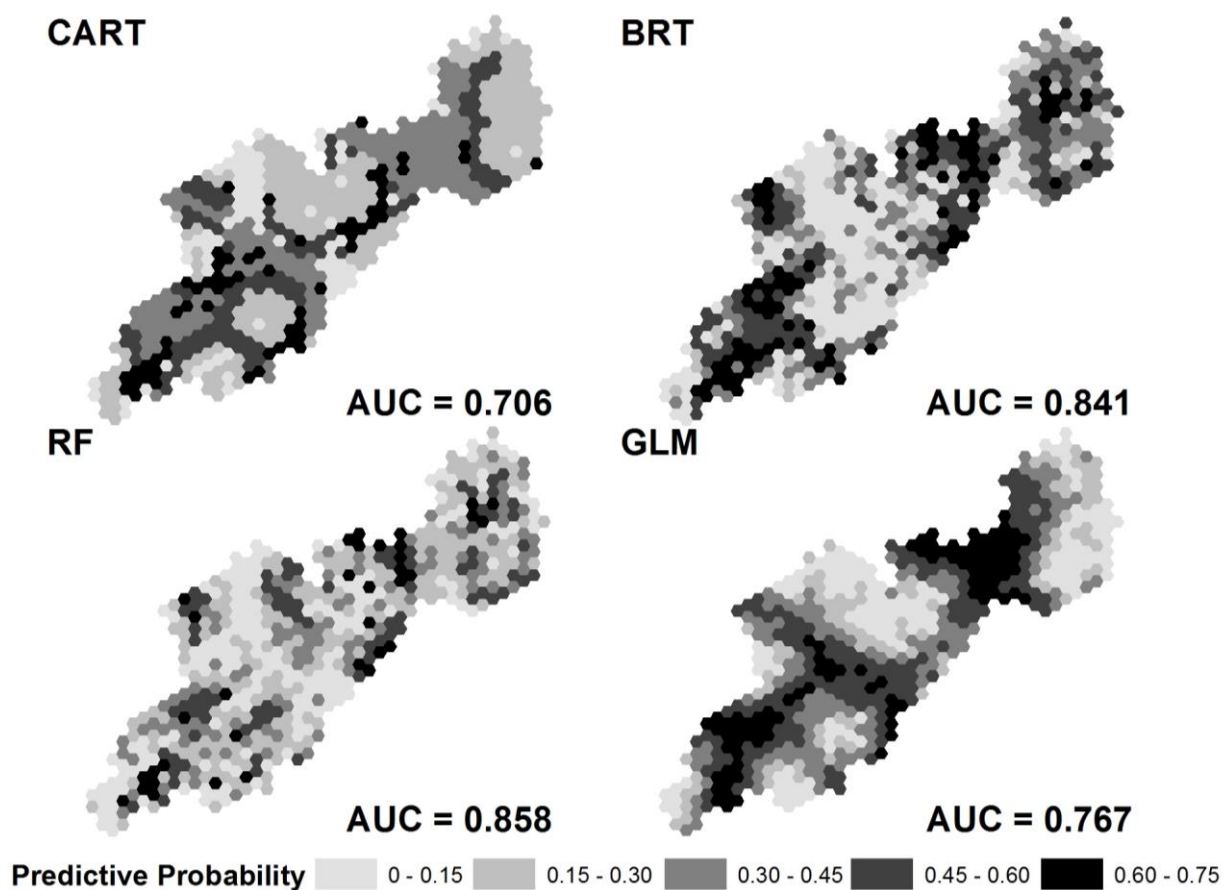


Figure 35 – Predictive probability of occurrence for red deer in the *Lousã* NHA according to the models obtained for *Lombada-Culebra* hunting areas.

Discussion

This study shows that the proposed methodologies are able to predict red deer probability of occurrence with high accuracy identifying those environmental predictors that determine the presence/absence of the species. The predictors deemed important by the models are partially in agreement. All the models has identified the distance to urban areas and distance to shrubs as the most important variables in predicting red deer occurrence. It is clear from the maps and statistics presented (see Figure 36) that the predictive performance of ensemble methods as BRT and RF was quite similar, being more effective than single classification methods and GLM. These ensemble algorithms based on individual trees can be viewed as “black boxes” however, as showed in this study, they are easily interpreted allowing the development of highly accurate predictions. The main characteristics of ecological/environmental matrices are the high dimensionality and collinearity which lead us to considerably differences in splitting rules among individual fitted trees. This model instability can be overcome through the application of bootstrap-resampling and averaging methods as boosted regression trees (boosting-based method) and random forests (bagging-based method). Considering our models insights and other studies where these approaches were compared we give more relevance to RF and BRT methods (Svetnik *et al.* 2003; Prasad *et al.* 2006). Understanding in advance why RF and BRT perform substantially better than other methods is difficult, however, non-linear relationships, strong correlations and additive effects among variables are possible reasons for the outperformance of ensemble approaches (Cutler *et al.* 2007). As reported by Prasad *et al.* (2006), RF high performance is related with a smoother response surface (see Figure 36) in that predictive probability gradually increases from lower to higher values without skipping classes. Further, the presence of modal effects (see partial dependence plots in Figure 26) could contribute for the lower performance of CT and GLM. Even though these partial dependency plots are not a perfect representation of the variables effects in red deer occurrence (Elith *et al.* 2008), namely if there are predictors highly correlated, they provide useful information for interpretation (Friedman 2001). In sparsely sampling, BRT fitted functions can exhibit a noisy behaviour but model performance was not affected (Elith *et al.* 2008).

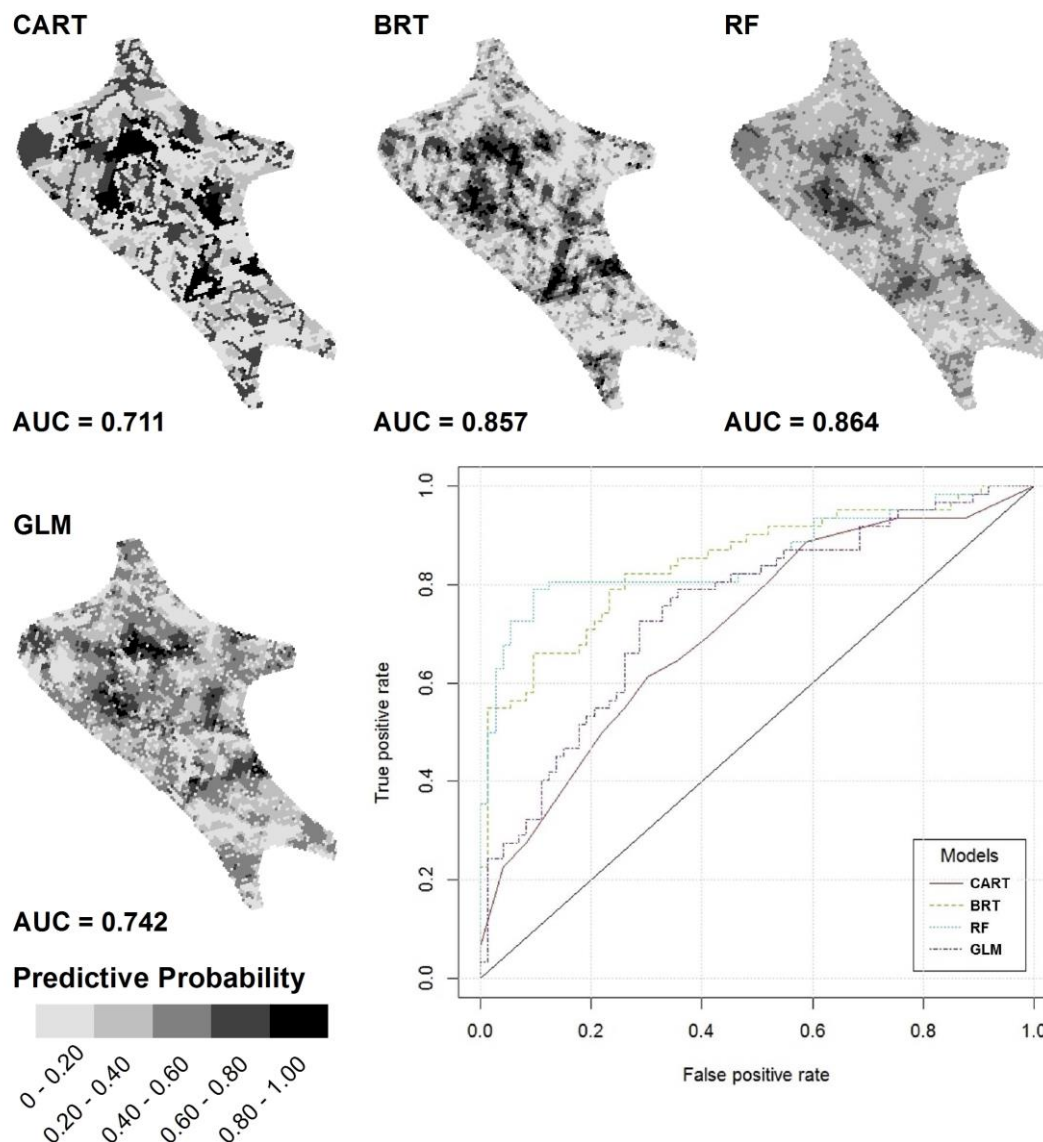


Figure 36 – The probabilities achieved were mapped and can be interpreted as an index of habitat suitability (Aspinall & Veitch 1993). A smoother response surface provided by RF is one of the main reasons for the model high performance.

From an ecological point of view, our results are consistent with the known ecology of red deer in the study area and are in agreement with previous studies carried out in the same region (Carvalho *et al.* 2012; Torres *et al.* 2012). Our study shows that red deer occurs far from disturbance areas (e.g. villages and human settlements), agricultural fields and near of shrubs and forest patches. Red deer often occurs in areas with gentle slopes preferentially exposed to south. Remaining variables exhibit modal effects, which makes it difficult to define patterns of selection. Torres *et al.* (2012) reported that high cover of

shrubs is positively related with red deer presence. In fact, some studies demonstrated a red deer preference for shrub species like *Pterospartum tridentatum*, *Cistus ladanifer*, *Halimium lasianthum* subsp. *alyssoides*, *Rubus ulmifolius* and *Erica* sp. (Alvarez & Ramos 1991; Ferreira 1998). Transition areas (e.g. ecotones) assume a great importance for red deer to cope with seasonal changes in resource availability (Putman & Flueck 2011). Red deer in our study area preferentially occurs far from agricultural fields, which are in disagreement with the described by Myrnerud *et al.* (2002). However, it is important to mention that such areas provide an important food resource especially during the winter. This fact can be an adaptation to contrasting seasonality, human disturbance or a hiding behaviour to reduce predation risk.

Our results proved that all four modelling frameworks are sufficiently reliable (AUC > 0.7) to be used in the prediction of species distribution. The ranking of variables according to their order of selection are positively correlated among methods, which indicate a partial agreement in the variable importance (Table 13).

The RF had the higher discriminative performance among all the models showing a slightly better AUC and HL values than BRT. RF increases classification reliability of CT aggregating the results of many individual trees, by averaging or voting, in one final accurate model (Breiman 2001). Two user-defined parameters, (i) number of trees and (ii) number of selected variables per node, should be optimized in order to reduce the OOB error. By restricting the number of explanatory variables available per each node ensures that correlation among fitted trees is small which improves the model performance. Here, considering the square root of the number of variables randomly selected per each node and a series of 10 ensemble sizes (minimum of 5 trees and maximum of 5000), the best model performance was achieved at 1000 trees (OOB error = 20.73; AUC = 0.864). By adding more trees to the ensemble, the AUC value and OOB error did not decrease nor increase.

The best BRT model was achieved using a *lr* of 0.005, a *tc* of 5 and a *bc* of 0.5. These tuning parameters have increased the model accuracy and the reliable selection of relevant variables. Through forward stagewise fitting and model averaging, BRT meta-algorithm reduces the bias and variance, respectively. RF cannot achieve bias reduction because the classification trees which form the ensemble are fitted in the same way (Elith *et al.* 2008). According to the same authors, the BRT algorithm settings need to be carefully selected when dealing with few records. CART models are intuitive and easily interpret providing useful insights of species-environment relationships, however, when

handling with sharp discontinuities of variables distributions, algorithm may have produced unstable decision trees. Moreover, when comparing with ensemble methods, CART tends to promote overfit due to the selection of tree with more splits, the identification of important interactions are compromised by strong correlations among variables and there is no guarantee that the final tree achieved represents the optimal model (Olden *et al.* 2008). Tree pruning and optimal tree selection are procedures that can reduce the error threshold. The overcoming of CART drawbacks is provided by the ensemble methods discussed above. The GLM stepwise procedure provides a statistical summary of the relationships between the explanatory variables and the response although the ecological interpretation is harder when compared with the unified view of variable effects offered by tree graphs. The GLM framework is particularly useful for testing hypothesis about the shape of response functions and variables significance (Franklin 1998). However, in some cases, the plotting of response surfaces is the only way to extract ecological meaning from the coefficients (Vayssières *et al.* 2000).

Despite the widespread use of CART in numerous ecological applications: development of habitat models for endangered (O'Connor *et al.* 1996; Anderson *et al.* 2000; Usio 2007) and game species (Debeljak *et al.* 2001), forecast species distributions shifting in response to climatic warming (Iverson & Prasad 1998), study of the relationship between forest fires and habitat structure (Rollins *et al.* 2004), among others, we consider that classification and regression tree models *per se* are not a reliable method to quantify the variable importance in SDMs. They provide a preliminary view of species-environment relationships and the identification of interaction terms but their instability does not allow the drawing of definitive conclusions. Ensemble classification methods as RF and BRT are starting to be applied in ecology (see De'ath 2007; Peters *et al.* 2007; Elith *et al.* 2008) and the number of comparisons are few (although see Prasad *et al.* 2006). Here, we found that the combination of the strengths of different classification methods provide useful and complementary insights in the study of species distributions. In fact, species' habitat preferences and biotic/abiotic drivers are tightly linked which requires an ensemble approach to identify the main interactions and describe the dominant patterns.

To our knowledge, this is the first study testing the spatial transferability of classification methods. The transference of models among study areas is one of the main challenges of predictive SDMs. We have tested model transferability using areas and study species whose distribution and density is well documented. Here, availability of presence records related to red deer occurrence for projected area and prior selection of

variables in order to guarantee the environmental similarity proved that this transference can be done with confidence. The descriptive ability of models outcomes is a key point when transferred or applied into a broad extent (Arntzen 2006). RF model scored the highest predictive performance in the both study areas however, the occurrence probabilities exhibit a diffuse pattern when projected outside from original scope, thus underestimating the occurrence probability of the species. Randin *et al.* (2006) argued that environmental differences between study regions and species-specific biotic factors are possible reasons for asymmetrical transferability. Indeed, there are environmental and biotic dissimilarities between *Lombada-Culebra* and *Lousã* hunting areas (i) the extent of *Lousã* area is significantly smaller, (ii) *Lombada-Culebra* is characterized by environmental combinations absent in the *Lousã* and (iii) red deer populations in Northwestern Iberian Peninsula are predated by wolf which can has important effects in habitat use. The use of background data instead of true absences can also influence the extrapolation success. Further, the removal of some variables to achieve an approximation of environmental similarity can reduce the models performance.

Notwithstanding the existence of some drawbacks, classification methods could be confidently applied in ecological distribution modelling outlining the main factors which affect the species distribution, occurrence and abundance. This study provides an effectiveness assessment of classification methods in the context of SDMs prompting the use of alternative and promising methods for interpretation and prediction.

References

- Acevedo P., Jiménez-Valverde A., Lobo J.M. & Real R. (2012). Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*. 39: 1383–1390.
- Akaike H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 19: 716-723.
- Alves J., Silva A.A., Soares A.M.V.M & Fonseca C. (2013). Pellet group count methods to estimate red deer densities: Precision, potential accuracy and efficiency. *Mammalian Biology*. 78: 134–141.
- Alvarez G. & Ramos J. (1991). Estrategias alimentarias del ciervo (*Cervus elaphus* L.) en Montes de Toledo. Doñana, *Acta Vertebrata*. 18: 63–99.
- Anderson M.C., Watts J.M., Freilich J.E., Yool S.R., Wakefield G.I., McCauley J.F. & Fahnestock P.B. (2000). Regression-tree modelling of desert tortoise habitat in the central Mojave desert. *Ecological Applications*. 10(3): 890–900.
- Araújo M.B., Thuiller W., Williams P.H. & Reginster I. (2005a). Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography*. 14: 17–30.
- Araújo M.B., Whittaker R.J., Ladle R. & Erhard M. (2005b). Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*. 14: 529–538.
- Araujo M.B., Thuiller W. & Pearson R.G. (2007). Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*. 33: 1712–1728.
- Araújo M.B. & New M. (2007). Ensemble forecasting of species distribution. *Trends in Ecology and Evolution*. 22: 42–47.

Arntzen J.W. (2006). From descriptive to predictive distribution models: a working example with Iberian amphibians and reptiles. *Frontiers in Zoology*. 3: 8.

Aspinall R. & Veitch N. (1993). Habitat mapping from satellite imagery and wildlife survey data using a Bayesian modelling procedure in GIS. *Photogrammetric Engineering and Remote Sensing*. 59: 537–543.

Austin M. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*. 200: 1–19.

Barbosa A.M., Real R. & Vargas J.M. (2008). Transferability of environmental favourability models in geographic space: The case of the Iberian desman (*Galemys pyrenaicus*) in Portugal and Spain. *Ecological Modelling*. 220: 747–754.

Breiman L., Friedman J.H., Olshen R.A. & Stone C.J. (1984). Classification and regression trees. Wadsworth International Group. 358pp.

Breiman L. (2001). Statistical modelling: the two cultures. *Statistical Science*. 16: 199–215.

Carvalho J., Martins L., Silva J.P., Santos J., Torres R.T. & Fonseca C. (2012). Habitat suitability model for red deer (*Cervus elaphus* Linnaeus, 1758): spatial multi-criteria analysis with GIS application. *Galemys*. 24: 47–56.

Cianfrani C., Maiorano L., Loy A., Kranz A., Lehmann A., Maggini R. & Guisan A. (*in press*). There and back again? Combining habitat suitability modelling and connectivity analyses to assess a potential return of the otter to Switzerland. *Animal Conservation*.

Clark L.A. & Pregibon D. (1992). Tree-based models. Pp. 377–420. In: Chambers J.M. & Hastie T.J. (ed.). *Statistical Models in S*. Wadsworth and Brooks. 624pp.

Clausnitzer V., Kalkman V.J., Ram M., Collen B., Baillie J.E.M., Bedjanic M., Darwall W.R.T., Dijkstra K.D.B., Dow R., Hawking J., Karube H., Malikova E., Paulson D., Schütte K., Suhling F., Villanueva R.J., Ellenrieder N. & Wilson K. (2009). Odonata enter the biodiversity crisis debate: The first global assessment of an insect group. 142(8): 1864-1869.

Cutler D.R., Edwards T.C., Beard K.H., Cutler A., Hess K.T., Gibson J. & Lawler J.J. (2007). Random forests for classification in ecology. *Ecology*. 88(11): 2783–2792.

De'ath G. & Fabricius K.E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*. 81(11): 3178–3192.

De'ath G. (2007). Boosted trees for ecological modelling and prediction. *Ecology*. 88(1): 243–251.

Debelkak M., Dzeroski S., Jerina K., Kobler A. & Adamic M. (2001). Habitat suitability modelling for red deer (*Cervus elaphus* L.) in South-central Slovenia with classification trees. *Ecological Modelling*. 131: 321–300.

Elith J. *et al.* (2006). Novel methods improve predictions of species' distributions from occurrence data. *Ecography*. 29: 129–151.

Elith J., Leathwick J.R. & Hastie T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*. 77: 802–813.

Elith J. & Leathwick J.R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*. 40: 677–697.

Evans S.E., Murphy M.A., Holden Z.A. & Cushman S.A. (2011). Modelling species distribution and change using Random Forest. Pp.139-160. Em: Drew C.A., Wiersma Y.F. & Huettmann F. (eds.). *Predictive Species and Habitat Modelling in Landscape Ecology – Concepts and Applications*. Springer. 313pp.

- Ferreira S. (1998). Estudo da dieta de duas populações de veado (*Cervus elaphus* Linnaeus, 1758) em Portugal. First degree thesis in Biology, University of Coimbra, Portugal.
- Fielding A.H. & Bell J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*. 24: 38-49.
- Franklin J. (1998). Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*. 9: 733–748.
- Freund Y. & Schapire R.E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*. 55(1): 119–139.
- Friedman J., Hastie T. & Tibshirani R. (2000). Additive logistic regression: a statistical view of boosting. *Annals of Statistics*. 28: 337–407.
- Friedman J.H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*. 29: 1189–1232.
- Guisan A. & Zimmermann N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*. 135: 147–186.
- Guisan A. & Thuiller W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*. 8: 993-1009.
- Hastie T. & Tibshirani R. (1990). *Generalised Additive Models*. Chapman and Hall. 335pp.
- Iverson L.R. & Prasad A.M. (1998). Predicting abundance of 80 tree species following climate change in the Eastern United States. *Ecological Monographs*. 68(4): 465–485.
- Jiménez-Valverde A. & Lobo J.M. (2006). Distribution determinants of endangered Iberian Spider *Macrothele calpeiana* (Araneae, Hexathelidae). *Environmental Entomology*. 35: 1491–1499.

Jiménez-Valverde A., Lira-Noriega A., Barve N., Maher S.P., Nakazawa Y., Papes M., Soberón J., Sukumaran J. & Peterson A.T. (2011). Dominant climate influences on North American birds. *Global Ecology and Biogeography*. 20: 114–118.

Jiménez-Valverde A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*. 21: 498–507.

Jiménez-Valverde A., Acevedo P., Márcia Barbosa A., Lobo J.M. & Real R. (2013). Discrimination capacity in species distribution models depends on the representativeness of the environmental domain. *Global Ecology and Biogeography*. 22: 508-516.

Legendre P. & Legendre L. (1998). *Numerical ecology*. 2nd English edn. Elsevier Science. 306pp.

Liaw A. & Wiener M. (2012). Package “randomForest”: Breiman and Cutler’s random forests for classification and regression. URL: <http://cran.r-project.org/web/packages/randomForest/randomForest.pdf>.

Lobo J.M., Jiménez-Valverde A. & Hortal J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*. 33: 103–114.

Marmion M., Parviainen M., Luoto M., Heikkinen R.K. & Thuiller W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*. 15: 59–69.

McCullagh P. & Nelder J.A. (1989). *Generalized Linear Models*. Chapman and Hall. 511pp.

Milborrow S. (2012). Package “rpart.plot”: Plot rpart models. Extends plot.rpart and text.rpart in the rpart package. URL: <http://cran.r-project.org/web/packages/rpart.plot/rpart.plot.pdf>.

Moreira L.M., Rosa J.L., Lourenço J., Barroso I. & Pimenta V. (1997). *Projecto Lobo. Relatório de Progressão 1996* (Cofinanciado pela U. E. - Programa LIFE). Parque Natural de Montesinho, Bragança, Portugal. 61 pp.

Mysterud A., Langvatn R., Yoccoz N.G. & Stenseth N.C. (2002). Large-scale habitat variability, delayed density effects and red deer populations in Norway. *Journal of Animal Ecology*. 71: 569–580.

O'Connor R.J., Jones M.T., White D., Hunsaker C., Loveland T., Jones B. & Preston E. (1996). Spatial partitioning of environmental correlates of avian biodiversity in the conterminous United States. *Biodiversity Letters* 3(3): 97–110.

Olden J.D., Lawler J.J. & Poff N.L. (2008). Machine learning methods without tears: a primer for ecologists. *Quarterly Review of Biology*. 83: 171–193.

Pearce J. & Ferrier S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*. 133: 225-245.

Peters J., De Baets B., Verhoest N.E.C., Samson R., Degroeve S., De Becker P. & Huybrechts W. (2007). Random forests as a tool for ecohydrological distribution modelling. *Ecological Modelling*. 207: 304–318.

Phillips S.J., Dudík M., Elith J., Graham C.H., Lehmann A., Leathwick J. & Ferrier S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Application*. 19: 181–197.

Phillips S.J. & Elith J. (2011). Logistic methods for resource selection functions and presence-only species distribution models, AAAI (Association for the Advancement of Artificial Intelligence). San Francisco. USA. 1384–1389.

Post E. & Forchhammer M.C. (2002). Synchronization of animal population dynamics by large-scale climate. *Nature*. 420: 168-171.

Prasad A.M., Iverson L.R. & Liaw A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*. 9: 181–199.

Putman R. & Flueck W.T. (2011). Intraspecific variation in biology and ecology of deer: magnitude and causation. *Animal Production and Science*. 51(4): 277-291.

Randin C.F., Dirnböck T., Dullinger S., Zimmermann N.E., Zappa M. & Guisan A. (2006). Are niche-based species distribution models transferable in space? *Journal of Biogeography*. 33: 1689–1703.

Real R., Márquez A.L., Estrada A., Munoz A.R. & Vargas J.M. (2008). Modelling chorotypes of invasive vertebrates in mainland Spain. *Diversity and Distributions*. 14: 364–373.

Ridgeway G. (2013). Package “gbm”: Generalized Boosted Regression Models. URL: <http://cran.r-project.org/web/packages/gbm/gbm.pdf>.

Rollins M.G., Keane R.E. & Parsons R.A. (2004). Mapping fuels and fire regimes using sensing, ecosystem simulation, and gradient modelling. *Ecological Applications*. 14(1): 75–95.

Sánchez-Fernández D., Lobo J.M. & Hernández-Manrique O.L. (2011). Species distribution models that do not incorporate global data misrepresent potential distributions: a case study using Iberian diving beetles. *Diversity and Distributions*. 17: 163–171.

Schadt S., Revilla E., Wiegand T., Knauer F., Kaczensky P., Breitenmoser U., Bufka L., Cerveny J., Koubek P., Huber T., Stanisa C. & Trepl L., 2002. Assessing the suitability of central European landscapes for the reintroduction of Eurasian lynx. *Journal of Applied Ecology*. 39: 189–203.

Schröder B. & Richter O. (1999). Are habitat models transferable in space and time? *Zeitschrift für Ökologie und Naturschutz*. 8: 195–205.

Strobl C., Malley J. & Tutz G. (2009). An introduction to recursive partitioning: rational, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*. 14(4): 323–348.

Svetnik V., Liaw A., Tong C., Culberson J.C., Sheridan R.P., Feuston B.P. (2003). Random Forests: a classification and regression tool for compound classification and QSAR modelling. *Journal of Chemical Information and Modelling*. 43(6):1947–1958.

Therneau T., Atkinson B. & Ripley B. (2013). Package “rpart”: Recursive partitioning and regression trees. URL: <http://cran.r-project.org/web/packages/rpart/rpart.pdf>.

Thomas J.A. & Bovee K.D. (1993). Application and testing of a procedure to evaluate transferability of habitat suitability criteria. *Regulated Rivers – Research and Management*. 8: 285–294.

Torres R.T., Virgós E., Santos J., Linnell J.D.C. & Fonseca C. (2012). Habitat use by sympatric red and roe deer in a Mediterranean ecosystem. *Animal Biology*. 62: 351 – 366.

Tsoar A., Allouche O., Steinitz O., Rotem D. & Kadmon R. (2007). A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions*. 13: 397–405.

Usio N. (2007). Endangered crayfish in northern Japan: distribution, abundance and microhabitat specificity in relation to stream and riparian environment. *Biological Conservation* 134(4): 517–526.

Vayssières M.P., Plant R.E. & Allen-Diaz B.H. (2000). Classification Trees: An Alternative Non-Parametric Approach for Predicting Species Distributions. *Journal of Vegetation Science*. 11: 679–694.

Vingada J., Fonseca C., Cancela J., Ferreira J. & Eira C. (2010). Ungulates and their management in Portugal. Pp. 392–418. In: Apollonio M., Andersen R. & Putman R. (eds). *European Ungulates and their Management in the 21st Century*. Cambridge University Press.

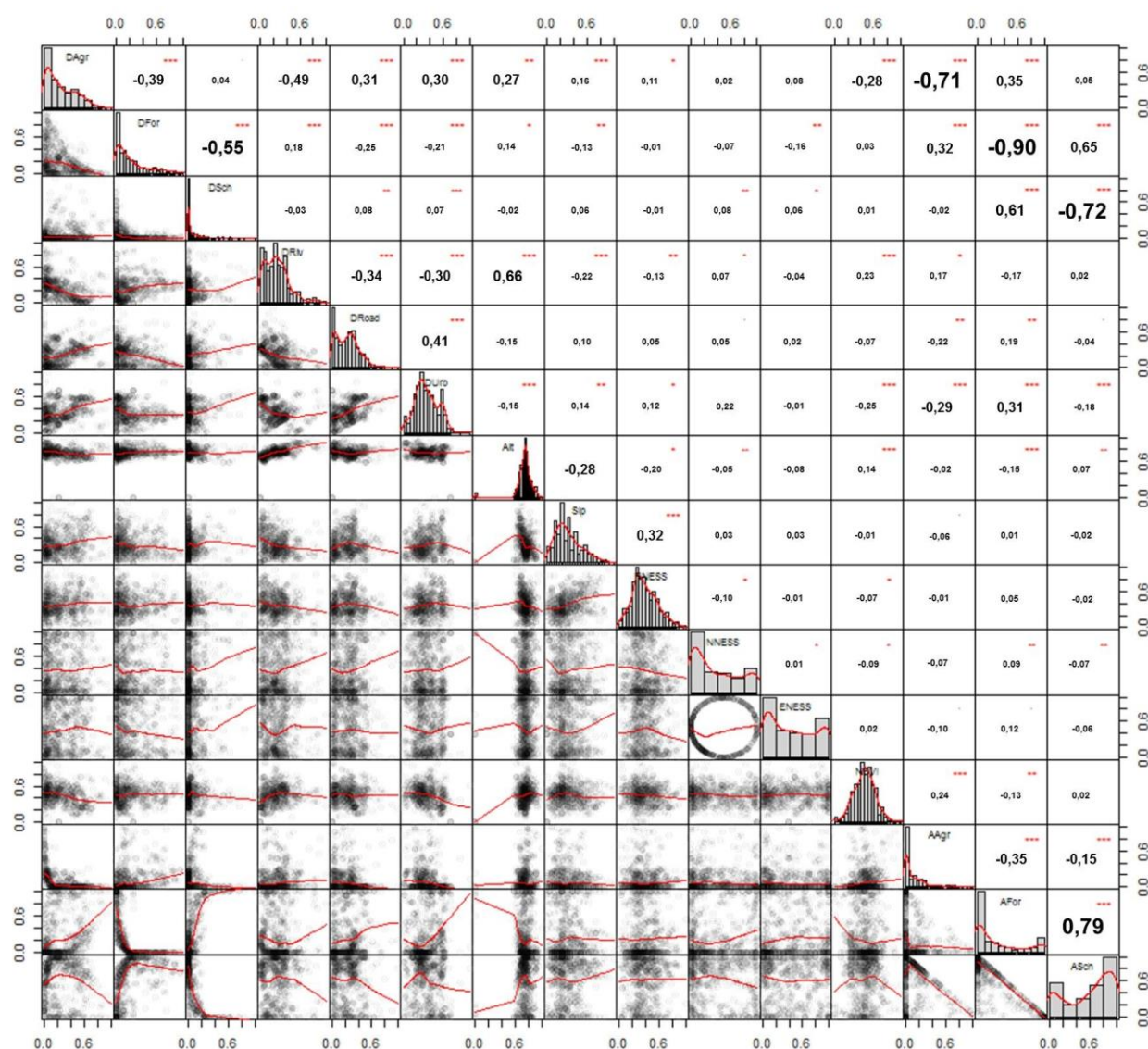
Warton D.I. & Shepherd L.C. (2010). Poisson point process models solve the “pseudo-absence problem” for presence-only data in ecology. *The Annals of Applied Statistics*. 3: 1383–1402.

Williams K.J., Belbin L., Austin M.P., Stein J.L. & Ferrier S. (2012). Which environmental variables should I use in my biodiversity model? *International Journal of Geographical Information Science*. 26(11): 2009-2047.

Zimmermann NE, Edwards TC, Graham CG, Pearman PB & Svenning JC (2010). New trends in species distribution modelling. *Ecography*. 33: 985-989.

Zuur A.F., Ieno E.N. & Elphick C.S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology & Evolution*. 1: 3-14.

Appendix I – Spearman correlation matrix representing the interactions among environmental variables.



Capítulo 6

Sumário dos principais resultados

Conclusões gerais

Principais resultados e conclusões gerais

O presente estudo teve como principal objetivo implementar uma abordagem holística no estudo e monitorização de populações silvestres. Procedeu-se à caracterização demográfica das populações mediante a determinação de vários parâmetros essenciais à sua correta gestão e exploração. Os principais resultados são sumariamente apresentados em seguida.

- A ZCN da Lombada e RRC *Sierra de la Culebra* registam um equilíbrio da estrutura populacional ($r_s = 0,78$ e $p_r = 0,36$) sendo que as discrepâncias observadas no que concerne ao rácio sexual são consideradas naturais em populações silvestres. Para o Inverno de 2009 foi estimada uma densidade populacional de 3,26 veados/100ha (IC95% = 2,27 – 4,70);
- Na ZCN da Lousã ($r_s = 0,32$ e $p_r = 0,40$) e ZCT da Cubeira ($r_s = 0,28$ e $p_r = 0,42$) o desequilíbrio é mais evidente refletindo o tipo de gestão cinegética a que estão sujeitas e o período em que foi realizada a amostragem. As densidades da espécie para o período Invernal foram estimadas em 5,52 (IC95% = 2,90 – 10,40) para a Lousã e 23,49 (IC95% = 15,83 – 34,85) para a Cubeira.

Relativamente às unidades sociais, estas variam ao longo do ano em número e composição. Da época de acasalamento que se desenrola no período outonal para o período fenológico seguinte, observa-se uma diminuição do número de grupos mistos e um incremento no tamanho médio dos grupos.

- Para a ZCN da Lombada e RRC *Sierra de la Culebra*, na época da brama, o tamanho médio dos grupos foi calculado em 2,25 (IC95% = 2,11 – 2,40). Da brama para a época seguinte registou-se um aumento de aproximadamente 50% do tamanho dos grupos cifrando-se em 3,72 (IC95% = 3,33 – 4,18);
- Na ZCN da Lousã observou-se a mesma tendência caracterizada por grupos com tamanho médio de 2,32 (IC95% = 2,09 – 2,59) na brama e 3,01 (IC95% = 2,64 – 3,50) para a período pós-brama;

- Na ZCT da Cubeira, o tamanho médio registado na brama foi de 2,79 (IC95% = 2,49 – 3,13) e de 4,03 (IC95% = 3,40 – 4,85) para o período pós-brama.

O capítulo 4 constitui uma abordagem exploratória testando a aplicabilidade de um método heurístico na determinação de quais as variáveis mais decisivas à ocorrência da espécie estudada, o veado. O carácter exploratório deste capítulo justificou o aprofundamento da temática no capítulo 5 através do aumento da abrangência espacial do estudo, do número de registos de presença e do número de variáveis avaliadas. Foram testadas um conjunto de cinco metodologias (uma estatística (*GLM*), uma heurística (*AHP*) e três de *machine learning* (*CT*, *RF* e *BRT*)) e avaliada a sua capacidade preditiva.

Do capítulo 4, mediante a utilização de seis variáveis explicativas (ver capítulo 4) retiraram-se as seguintes conclusões:

- A divisão da ZCN da Lombada em três classes de adequabilidade determinou que 58% do total da zona de caça apresenta uma adequabilidade elevada, 27% uma adequabilidade moderada e os restantes 15% representam as áreas de reduzida adequabilidade;
- O valor AUC atribuiu uma capacidade discriminatória ligeiramente superior ao método heurístico (AUC = 0,72) relativamente à regressão logística (AUC = 0,69);
- Apesar da subjetividade inerente aos métodos heurísticos, este estudo provou que, desde que devidamente fundamentados e validados, a sua aplicabilidade em modelos de distribuição de espécies é fiável.

O capítulo 5 constitui uma análise *ensemble* mais abrangente tendo sido consideradas 15 variáveis explicativas, duas áreas de estudo e um total de 691 registos de presença da espécie (539 para a população do Noroeste peninsular e 152 para a ZCN da Lousã). Os principais resultados e conclusões foram:

- A probabilidade de ocorrência do veado aumenta à medida que a distância a aglomerados populacionais e campos agrícolas aumenta. Por sua vez, a diminuição da distância a áreas de mato, manchas florestais e rios traduz-se num

incremento nessa mesma probabilidade. A espécie está igualmente associada a áreas de declives suaves preferencialmente expostas a sul;

- Dos modelos testados, o *RF* apresentou o maior poder discriminatório sendo igualmente o mais calibrado seguido pelo *BRT*, *GLM* e *CART*;
- Os modelos desenvolvidos na ZCN da Lombada e RRC *Sierra de la Culebra* foram transferidos e projetados na ZCN da Lousã sendo validados com um total de 152 registos de presença da espécie. A ordem da *performance* preditiva manteve-se a mesma: *RF* (AUC = 0,858), *BRT* (AUC = 0,841), *GLM* (AUC = 0,767) e *CART* (AUC = 0,706).

Os resultados obtidos na presente dissertação permitem uma caracterização pormenorizada de três populações de veado distintas, constituindo informação relevante no que concerne à sua ecologia e fatores que determinam a sua distribuição.

Os parâmetros demográficos são elucidativos do equilíbrio das populações. As populações da ZCN da Lousã e ZCT da Cubeira apresentam, de acordo com os resultados obtidos, estruturas populacionais desequilibradas o que sugere, dependendo dos objetivos de gestão, uma necessidade de correção dos seus efetivos. No que respeita ao uso do *habitat*, destacam-se as áreas de matos desenvolvidos e as manchas florestais. Estas providenciam alimento, abrigo e refúgio à espécie adquirindo maior ou menor importância de acordo com o grau de perturbação (e.g. antropogénica, predação) a que a mesma está sujeita. A conservação e fomento destes *habitats* aliados a uma estrutura populacional equilibrada são dois dos fatores chave para o sucesso da gestão do veado.

A título conclusivo, realça-se a importância de uma gestão integrada das populações silvestres mediante a sua caracterização demográfica, identificando as variáveis bióticas e abióticas mais importantes que determinam a sua distribuição e abundância. A multidisciplinariedade deste estudo constitui uma abordagem holística e metodológica que poderá e deverá ser aplicada a escalas mais abrangentes e a outras espécies tendo em vista uma exploração cinegética sustentável.